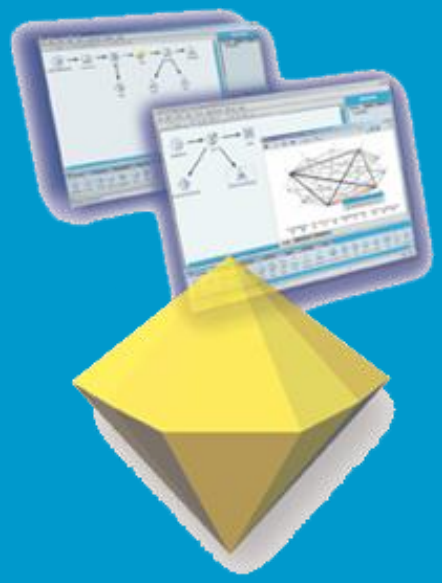
SAP BO Predictive Workbench

Voorspellende analyses



Afstudeerscriptie

accenture

High performance. Delivered.

Datum:	10 januari 2011
Naam:	Sascha Maas
Studentnummer:	2098638
Afstudeerbedrijf:	Accenture
Bedrijfsbegeleiders:	Irina Mizhiritskaya Alexander Radin
Eerste afstudeerdocent:	Ko Vleugel
Tweede afstudeerdocent:	Mart van Hulten
Versie:	1.0
Status:	Final

Opdrachtgever

Bedrijfsnaam:	Accenture
Bezoekadres:	Hambakenwetering 5K, 2e verdieping
Postcode en plaats:	5231 DD Den Bosch
Telefoon:	+31 20 593 9900
Internet:	www.accenture.com
Afdeling:	ATS (BI)

Begeleiders vanuit organisatie

Naam:	Irina Mizhiritskaya
Afdeling:	Solutions/BI
Telefoonnummer:	+31 20 5729297
E-mail adres:	irina.mizhiritskaya@accenture.com

Naam:	Alexander Radin
Afdeling:	Solutions/BI
Telefoonnummer:	+31 20 5729353
E-mail adres:	alexander.radin@accenture.com

Begeleider vanuit opleiding

Naam:	Ko Vleugel
Organisatie onderdeel:	Fontys Hogescholen ICT
Gebouw:	Eindhoven, Rachelsmolen 1 (R1)
Kamer:	4.07
Telefoonnummer:	0877 870988
E-mail adres stage:	k.vleugel@fontys.nl

Afstudeerder

Naam:	Sascha Maas
Adres:	Weisshorn 175
Postcode:	5624 NW Eindhoven
Telefoonnummer:	06-42454421
E-mail adres stage:	sascha.maas@accenture.com
E-mail adres privé:	maas.sascha@gmail.com

Documentenhistorie

Revisies

Versie	Status	Datum	Wijzigingen
0.1	Tussenrapportage	01-10-2010	Initiële uitgave
0.2	Concept	03-01-2011	Uitwerking
0.3	Concept	04-01-2011	Feedback stagebegeleider Fontys verwerkt
0.4	Concept	06-01-2011	Feedback bedrijfsbegeleiders Accenture verwerkt
1.0	Final	07-01-2011	Scriptie afgerond

Distributie

Dit document is verstuurd aan:

Versie	Datum	Naam	Functie
0.1	01-10-2010	Stage- en afstudeerbureau	Secretariaat
		Ko Vleugel	1e afstudeerdocent
		Mart van Hulten	2e afstudeerdocent
0.2	03-01-2011	Irina Mizhiritskaya	Bedrijfsbegeleider
		Alexander Radin	Bedrijfsbegeleider
		Ko Vleugel	1e afstudeerdocent
1.0	10-01-2011	Stage- en afstudeerbureau	Secretariaat

Voorwoord

Deze scriptie is geschreven naar aanleiding van mijn afstudeeropdracht: 'SAP BusinessObjects Predictive Workbench' die is uitgevoerd bij Accenture. Dit document geeft inzicht in de aanpak van het project, het onderzoek, de uitvoering en de resultaten.

In het project staat het onderzoek naar de mogelijkheden van de tool SAP BusinessObjects Predictive Workbench, waarmee onder andere voorspellende analyses mee gemaakt kunnen worden, centraal. Het onderwerp van deze afstudeeropdracht sprak me heel erg aan. Het leek me erg interessant om (voorspellende) analyses te maken waar bedrijven vervolgens beslissingen op kunnen baseren waardoor doelgerichter gewerkt kan worden. Hiernaast vond ik het wiskundige deel van de opdracht ook erg leuk om me in te verdiepen.
Tijdens mijn afstudeerstage heb ik gekeken wat de mogelijkheden zijn van de tool SAP BusinessObjects Predictive Workbench. Doel hierbij was om te bekijken of het voor Accenture interessant is om deze tool in de toekomst bij klanten in te gaan zetten.

Op deze plek wil ik van de gelegenheid gebruikmaken een aantal mensen, in willekeurige volgorde, te bedanken voor hun positieve bijdrage aan mijn afstudeerstage.
Ik wil graag mijn stagebegeleiders Alexander Radin en Irina Mizhiritskaya bedanken voor de ondersteuning die zij mij gedurende mijn afstudeerperiode hebben gegeven. Daarnaast wil ik ook graag Accenture bedanken dat ik mijn afstudeerstage heb kunnen lopen bij dit bedrijf.
Ook wil ik graag mijn stagedocent Ko Vleugel bedanken voor de ondersteuning vanuit Fontys Hogescholen.

Allen hebben een bijdrage geleverd en mij goed ondersteund tijdens het uitvoeren van dit project.

Den Bosch, januari 2011

Inhoudsopgave

1 INLEIDING 9

2 BESCHRIJVING ACCENTURE 10

2.1 GESCHIEDENIS 10
2.2 VESTIGINGEN 11
2.3 ORGANISATIESTRUCTUUR 11

3 DE OPDRACHT 12

3.1 ACHTERGROND 12
3.2 DOEL VAN HET PROJECT 12
3.3 OPDRACHTDEFINITIE 12

4 BUSINESS INTELLIGENCE 13

5 PREDICTIVE ANALYTICS 14

5.1 DATAMINING 14
5.2 BESCHRIJVENDE ANALYSES 14
5.3 VOORSPELLENDE ANALYSES 15

6 ANALYSEMETHODES 17

6.1 KLANT ANALYSE 17
6.2 FRAUDE 17

7 SAP BO PREDICTIVE WORKBENCH 18

7.1 SAP 18
7.2 BUSINESSOBJECTS 18
7.3 TOOL 19

8 BRANCHES 22

9 WERKZAAMHEDEN ONDERZOEK 23

9.1 ONDERZOEK TOOL 23
9.2 KOPPELINGEN 23

10 BUSINESS CASE 25

10.1 PROBLEEMDEFINITIE 25
10.2 OPLOSSINGEN 26
10.3 INVESTERINGSBEOORDELING 28
10.4 BELANGRIJKSTE RISICO'S 30

11 CONCLUSIE(S) EN AANBEVELING(EN) 31

EVALUATIE 32

LITERATUURLIJST 33

BIJLAGEN 34

Samenvatting

Deze afstudeeropdracht is uitgevoerd bij Accenture, een internationaal consultancy bedrijf dat diensten aanbiedt voor management, technologie en outsourcing.

Het doel van dit project is het inzichtelijk maken van de mogelijkheden van de tool SAP BO Predictive Workbench voor Accenture. SAP BO Predictive Workbench is een nieuwe geïntegreerde tool binnen SAP welke het voor bedrijfsanalisten mogelijk maakt voorspellende analyses te realiseren in een gebruiksvriendelijke en toegankelijke omgeving. Op basis van dit onderzoek kan Accenture vervolgens gaan bekijken of het interessant is om de tool in de toekomst te gaan implementeren bij klanten.

Predictive Analytics is de theorie achter de tool en hieronder wordt verstaan het intelligent en doelgericht gebruiken van beschikbare data. Predictive Analytics bestaat uit een aantal technieken van statistiek en datamining die actuele en historische feiten analyseren om voorspellingen te maken van toekomstige gebeurtenissen. Deze voorspellende technieken zijn uitermate geschikt als een organisatie bijvoorbeeld wil weten welke klanten een hoge kans hebben om over te stappen naar de concurrentie.

De tool is door de eenvoudige grafische interface relatief eenvoudig te gebruiken. Kennis van statistiek of analytics is hierbij echter wel noodzakelijk om de werkwijze van de analysemodellen te begrijpen en de uitkomst goed te interpreteren en weer te geven. De gemaakte analyses kunnen gevisualiseerd worden waardoor de uitkomst voor alle eindgebruikers op een snelle manier inzichtelijk wordt.

De krachtige oplossing van SAP BO Predictive Workbench kan worden gebruikt als ‘glazen bol’ om de organisatie alert te houden en de concurrentie één stap voor te blijven.
Voordelen die kunnen ontstaan met de tool SAP BO Predictive Workbench zijn bijvoorbeeld een hogere conversie van campagnes of het kostenefficiënt, gericht en effectief benaderen van klanten.

In dit project is gekeken naar de werkwijze van SAP BO Predictive Workbench. Vervolgens is een demo uitgewerkt waarbij churn (weglopen van klanten) in een organisatie centraal staat. Aan de hand van de tool zijn analyses gemaakt waardoor er meer inzicht ontstond in het klantgedrag en de mogelijke churn kon worden voorspeld. Door de kans op churn bij klanten te voorspellen, kan hierop proactief worden gereageerd door middel van bijvoorbeeld een gerichte campagne. Met het benaderen van deze specifieke klanten kan mogelijk een verhoging van de retentie (klantbehoud) worden gerealiseerd.

Naar aanleiding van het onderzoek is naar voren gekomen dat predictive analytics zeker in de toekomst heel belangrijk gaat worden voor bedrijven. Door de steeds zwaardere concurrentie zullen bedrijven het belang van predictive analytics waarschijnlijk in gaan zien waarop de aanschaf van deze software een logisch gevolg is.

De tool SAP BO Predictive Workbench zou zeer geschikt zijn om binnen bedrijven voorspellende analyses te realiseren. De tool biedt voor verschillende branches mogelijkheden waar veel voordeel mee kan worden behaald. Uiteindelijk kan een organisatie beslissingen gaan sturen en optimaliseren door het integreren van voorspellende technologie op de huidige situatie met als doel de winstgevendheid te maximaliseren.

Summary

This research project was completed spanning five months during graduation and is commissioned by Accenture, an international consultancy company which provide management, technology and outsourcing services.

The goal of this project is to get insight about the possibilities of the tool SAP BO Predictive Workbench. SAP BO Predictive Workbench is a new integrated tool within SAP which can be used by business analysts to realize predictive analyses in a user friendly and accessible environment. Based on the research project, Accenture would be able to decide if the tool is interesting and useful enough to implement on a client side.

Predictive Analytics is using the available data in an intelligent and targeted way. Predictive Analytics consists of a few statistical and data mining techniques which analyze actual and historical facts to make predictions of upcoming events. These predictive techniques are very useful when, for example, an organization wants to find out which clients could switch to a competitor.

The tool is because it's graphical interface quite easy to use. However, knowledge about statistic and analytics is needed to understand the method used by the analysis models and to show and interpreting the outcome the right way. The analyses made, can be visualized which will make the results understandable and clear for the end-user.

The powerful solution Predictive Workbench can be used as a crystal ball for an organization to stay alert and be one step ahead of the competitor.
Benefits of using the SAP BO Predictive Workbench tool could be a higher campaign conversion or approach customers in a cost-efficient, direct and efficient way.

During this project the methods of SAP BO Predictive Workbench has been discovered. A demo about churn (leaving customers) in an organization is realized. Using the tool, analyses are made which has led to more insight in customer behavior and the possible churn could be predicted. By predicting possible churn it is possible to react pro-active by using a specific campaign. With approaching these specific customers, higher customer retention should be realized.

This research project has made clear that predictive analytics is a very important subject in the near future. With the ever increasing competition, companies will probably see the importance of this kind of software, and the purchase of this software is a logical consequence.

The tool SAP BO Predictive Workbench could be very useful for organizations to realize predictive analyses. The tool provides many options for different industries where much benefit can be achieved. Finally, an organization could control and optimize decisions by integrating predictive technology to the current way of work in order to maximize profitability.

Verklarende woordenlijst

Analytics	Analytics maken gebruik van data in een specifieke toepassing of functioneel proces om context gerelateerde inzichten te verwerven, die vervolgens leiden tot uitvoerbare acties
Business Case	Een business case is een zakelijke afweging waarin een project beschreven wordt. In de business case worden de kosten tegen de baten afgewogen. Vaak wordt aan de hand van de business case besloten om wel of niet te starten en/of verder te gaan met een project.
Business Objects	SAP BusinessObjects omvat een groot aantal tools en toepassingen waarmee de prestaties van een onderneming kunnen worden verbeterd. SAP BO Predictive Workbench en Xcelsius zijn onderdelen van BusinessObjects.
BW	Een verzamelplaats waar gegevens uit operationele bedrijfssystemen op een eenvoudige wijze bij elkaar zijn gebracht voor het uitvoeren van analyses en selecties en het generen van managementinformatie.
Churn	Het begrip dat men de ene aansluiting (bijvoorbeeld telefoon) opzegt voor een nieuwe aansluiting bij een bijvoorbeeld een andere service-aanbieder.
KPI	Kritieke prestatie-indicatoren zijn variabelen om prestaties van ondernemingen te analyseren. Het is een managementinstrument.
Query	Met een query (Engels voor vraagstelling) wordt in de informatica een opdracht bedoeld die aan een database wordt gegeven om een bepaalde actie uit te voeren, die een selectie van gevonden records oplevert die voldoen aan de zoekcriteria.
SAP	De afkorting SAP staat voor: Systems, Applications and Products in Data processing. SAP is een wereldwijd opererend bedrijf dat zich specialiseert in bedrijfssoftware oplossingen.
Universe	Een semantische tussenlaag in SAP BusinessObjects waar de data wordt gemapped naar objecten met gebruik van business termen. Daardoor worden business gebruikers afgeschermd van de technische aspecten van dataopslag in de database. Universe is de basis voor het bouwen van rapporten.
Xcelsius	Xcelsius is een onderdeel van Business Objects en biedt gebruikers de mogelijkheid bedrijfsinformatie, op een visuele en interactieve manier te presenteren.

1 Inleiding

‘Wie in de toekomst kan kijken, heeft goud in handen.’

Op dit moment komen beslissingen die worden genomen in een bedrijf vaak voort uit rapporten met gebeurtenissen uit het heden of verleden. Rapporten maken is nuttig binnen het bedrijf, maar onvoldoende om vandaag beslissingen over morgen te kunnen nemen. Organisaties hebben de mogelijkheid nodig om de toekomst te beïnvloeden op het gebied van planning en zakelijke beslissingen.

De tool SAP BusinessObjects Predictive Workbench maakt het mogelijk om voorspellende analyses te maken. In essentie gaat het erom dat voorspellende modellen kijken naar een scenario uit het verleden of heden en daarop statistieken loslaten die proberen te voorspellen hoe het bedrijf het zou doen als het op dezelfde manier door zou gaan.

Het doel van mijn afstudeeropdracht is om te onderzoeken wat de mogelijkheden van SAP BO Predictive Workbench zijn met betrekking tot voorspellende analyses. Na dit te hebben onderzocht kan er worden gekeken of deze tool in de toekomst eventueel bij klanten van Accenture kan worden geïmplementeerd en wat hierbij komt kijken.

Als basis voor dit onderzoek is begonnen met het uitwerken van de theorie achter de tool SAP BO Predictive Workbench: ‘Predictive Analytics’. Na deze theorie te hebben uitgewerkt, konden de mogelijkheden van de tool worden ontdekt om uiteindelijk hiervan een demo te maken naar aanleiding van een business case die is opgesteld.

In hoofdstuk 2 is een beschrijving gegeven van het bedrijf Accenture waar deze afstudeeropdracht is uitgevoerd. In hoofdstuk 3 wordt de achtergrond, het doel en de definitie van de opdracht beschreven. In hoofdstuk 4 wordt kort ingegaan op de mogelijkheden van business intelligence. In hoofdstuk 5 is de theorie van predictive analytics uitgewerkt met een aantal voorbeelden van analysemodellen. In hoofdstuk 6 zijn vervolgens de analysemethodes klant- en fraude analyse uitgewerkt. Hoofdstuk 7 is volledig gewijd aan de mogelijkheden en de systemen rondom de tool SAP BO Predictive Workbench. In hoofdstuk 8 wordt kort ingegaan op het voordeel dat branches kunnen behalen met de tool. In hoofdstuk 9 staan de werkzaamheden van het onderzoek vermeld. Hoofdstuk 10 bestaat uit de business case die is opgesteld met daarin onder andere de reden voor het project en uiteindelijk de demo. Hoofdstuk 11 bestaat ten slotte uit de conclusie met de aanbevelingen.

2 Beschrijving Accenture[1]

Accenture is een internationaal consultancy bedrijf dat diensten aanbiedt voor consulting, technologie en outsourcing.

Consulting heeft als doel de bedrijfsresultaten te optimaliseren. Accenture heeft een diepgaande betrokkenheid bij een organisatie met strategische doelen en zorgt voor effectieve oplossingen. Hierbij kiest Accenture niet alleen voor een traditionele benadering, maar juist dankzij de breedte van de dienstverlening ook voor het ontwikkelen, implementeren en beheren van innovatieve IT gerelateerde oplossingen.

Bij technologie worden business oplossingen van onder meer SAP, Siebel, People Soft en Oracle ontwikkeld, getest en geïmplementeerd. Accenture vertaalt innovatieve technologie in innovatieve oplossingen. De wisselwerking tussen Technologie en Consulting is sterk en biedt zowel voor Accenture als voor opdrachtgevers interessante kansen en mogelijkheden.

Outsourcing diensten gaan van applicatiemanagement tot optimalisatie van bedrijfsprocessen en infrastructuur outsourcing. Het wordt ingezet als een onmisbaar instrument om de prestaties van klanten te verbeteren.

2.1 Geschiedenis

Accenture heeft een bewogen historie, die meer dan een halve eeuw teruggaat. Van de eerste dagen als pionier in een nieuwe wereld van informatietechnologie tot de positie vandaag de dag als marktleider op het gebied van Consulting, Technologie en Outsourcing.

2.1.1 De start

Het begon in 1913, toen de Amerikaan Arthur Edward Andersen samen met een partner besloot Arthur Andersen & Co op te richten. Tot aan zijn dood in 1947 was Andersen als senior partner nauw verbonden met de ontwikkeling en groei van zijn bedrijf.

In deze jaren 40 boden de recent uitgevonden computers de mogelijkheid voor bedrijven om hun productiviteit aanzienlijk te verhogen. Er waren echter maar weinig captains of industry die het volledig potentieel hiervan erkenden. Visionaire leiders van Arthur Andersen als Leonard Spacek, John Higgings en Joe Glickauf waren de uitzonderingen. Zij gaven de computers een plaats in de dagelijkse bedrijfsvoering en ontwikkelden zo een belangrijke rol voor consultants in de introductie van computertechnologie bij andere bedrijven.
Technologische innovatie ligt diep geworteld in de cultuur van Accenture.

Afbeelding 1 – Joe Glickauf

[1] Literatuurlijst: bron 12

2.1.2 Consulting & outsourcing

Klanten begonnen zich steeds meer te richten op het inzetten van computertechnologie om effectiever en efficiënter te gaan werken. Zodoende begon ook consulting binnen Arthur Andersen een steeds belangrijkere rol in te nemen en groeide het bedrijf uit tot een leidend consultancy bedrijf op wereldniveau.

In 1989 vormde een aantal partners van deze divisie onder de naam Andersen Consulting een nieuwe organisatie, waarbij de focus lag op Consulting en Technology Services. In 1991 kwam hier ook Outsourcing bij, toen Andersen Consulting haar eerste grote outsourcing project startte door accounting, finance en andere functies uit te gaan voeren. De wereldwijde verspreiding van technologie zorgde voor een competitieve druk op bedrijven in verschillende industrieën. Andersen Consulting begon met het aanbieden van geïntegreerde bedrijfsoplossingen op gebied van consulting, technology en outsourcing in verschillende industrieën.

2.1.3 Accent op de toekomst

In de jaren daarna ontstond een steeds grotere afstand tussen beide organisaties. Het mondde uiteindelijk uit in een definitieve scheiding in 2001. Andersen Consulting ging verder onder de naam Accenture, een samentrekking van 'Accent' on the 'Future'. Deze vondst van een Noorse medewerker symboliseerde op compacte wijze de drang om een internationale voorhoedespeler te zijn.

Afbeelding 2 – Think global, act local

Accenture heeft zich vanaf 2001 explosief ontwikkeld. Het is in Nederland tegen de economische ontwikkelingen in met gemiddeld meer dan 20 procent per jaar gegroeid. Accenture heeft een internationale geschiedenis en op dit moment een netwerk van vestigingen in 49 landen. Deze wereldwijde slagkracht combineert het met een solide plaatselijke verankering. 'Think global, act local' is een motto dat ook Accenture Nederland ver heeft gebracht.

2.2 Vestigingen

Accenture is een internationale organisatie met 110 kantoren in 49 landen. In bijlage B staan een aantal landen weergegeven die hiertoe behoren. De hoofdvestiging in Nederland is in Amsterdam. Hiernaast zijn er ook nog vestigingen in Den Bosch en Almere.

2.3 Organisatiestructuur

In bijlage C is het organisatieschema van Accenture weergegeven.
De opdracht zal worden uitgevoerd bij de cluster: 'SAP Solutions', onder de area: 'SAP BI & Finance'.

3 De opdracht

De afstudeeropdracht bestaat uit het onderzoeken van de mogelijkheden van voorspellende analyses en de toepassing hiervan in de tool SAP BO Predictive Workbench. SAP BO Predictive Workbench is een krachtige software voor geavanceerde voorspellende analyses. In de volgende paragraven zal eerst de achtergrond van de opdracht worden beschreven. Vervolgens zal in worden gegaan op het doel van het project en de precieze opdrachtdefinitie.

3.1 Achtergrond

Accenture heeft bij een groot aantal klanten business Intelligence oplossingen geïmplementeerd. Hierbij wordt doorgaans gebruik gemaakt van SAP ERP software. SAP is een geïntegreerd informatie- en besturingssysteem waarin bedrijfsmatige processen kunnen worden vastgelegd en beheerd.
Op dit moment maken bedrijven voornamelijk analyses die gericht zijn op het verleden. Voorspellende analyses, gericht op de toekomst, worden echter nog niet of nauwelijks gebruikt.

SAP heeft recentelijk een aantal software pakketten overgenomen waaronder BusinessObjects (BO). SAP BO Predictive Workbench is een onderdeel van BusinessObjects en een nieuwe geïntegreerde tool binnen SAP welke het voor bedrijfsanalisten mogelijk maakt voorspellende analyses te realiseren in een gebruiksvriendelijke en toegankelijke omgeving.

Accenture ziet potentie in deze nieuwe tool maar op dit moment wordt deze nog niet ingezet bij klanten. Hierbij bestaat nog onduidelijkheid over de specifieke bestaande problemen die middels deze tool opgelost kunnen worden. Tevens is het uiteindelijk te bereiken doel bij de klant niet helder. Onderzoek moet uitwijzen waar, wanneer en hoe deze tool in de praktijk kan worden ingezet.

3.2 Doel van het project

Het doel van dit project is het voor Accenture inzichtelijk maken van de mogelijkheden van de tool SAP BO Predictive Workbench. Op basis van dit onderzoek kan Accenture vervolgens gaan bekijken of het interessant is om de tool in de toekomst te gaan implementeren bij klanten.

De doelstellingen die voor dit project gedefinieerd zijn:

- Theorie Predictive Analytics beheersen;
- Werkwijze van SAP BO Predictive Workbench inzichtelijk maken;
- De tool SAP BO Predictive Workbench inhoudelijk beheersen;
    - Daarbij moet een oordeel gegeven kunnen worden over de uitvoering van verschillende typen analyses;
- SAP BO Predictive Workbench toepassen op de business case;
- Koppeling vanuit SAP BO Predictive Workbench maken met Business Objects;
- Resultaten mooier/beter visualiseren in Xcelsius.

3.3 Opdrachtdefinitie

Tijdens de opdracht zal onderzoek worden gedaan naar de tool SAP BO Predictive Workbench in samenwerking met Business Intelligence. Hierbij zal worden bekeken welke inspanningen nodig zijn voor het implementeren van de tool binnen organisaties en hoe deze het beste ingezet kan worden bij klanten van Accenture.

4 Business Intelligence[2]

Business intelligence (BI) kan worden omschreven als het proces van gegevens omzetten in informatie, wat vervolgens leidt tot kennis. Business intelligence heeft als doel een competitief voordeel te creëren en wordt als een waardevolle kerncompetentie beschouwd.

Organisaties hebben informatie nodig om goed te kunnen functioneren. Zonder betrouwbare en relevante informatie verliezen organisaties het zicht op de interne gang van zaken en externe ontwikkelingen, raken uit koers of balans en verspillen tijd en geld door de verkeerde dingen te doen of de goede zaken na te laten. Dit kan worden voorkomen door de beschikbare informatie goed toegankelijk te maken en deze ook daadwerkelijk op een juiste manier te laten gebruiken.
Een intelligente organisatie stelt zich open voor relevante signalen van binnen en buiten de organisatie en vertaalt deze op slimme wijze naar informatie en kennis die aanzet tot actie, wat kan leiden tot effectiever gedrag en meer doordachte handelswijzen. Met effectieve business intelligence kunnen organisaties veel tijd en gericht aandacht besteden aan het beslissen, organiseren en innoveren zelf. De activiteiten en bedrijfsprocessen van de organisatie kunnen soepeler gaan lopen en worden steeds beter op elkaar afgestemd.

In afbeelding 3 is een voorbeeld gegeven van het voordeel dat kan worden behaald door middel van business intelligence. In deze afbeelding is het concurrentievoordeel uitgezet tegenover de intelligentie van een organisatie. Door een stap hoger te gaan in dit schema wordt de organisatie steeds intelligenter waardoor er meer concurrentievoordeel kan worden gerealiseerd.

Afbeelding 3 - Business Intelligence

Wanneer geen analytics wordt gebruikt, wordt voornamelijk gewerkt met rapporten, query's en alerts (waarschuwingen). Hierbij wordt pas gereageerd op problemen nadat ze opgetreden zijn. Door gebruik te maken van analytics wordt er gekeken waarom dingen gebeuren, welke trends hierbij meespelen en wat in de toekomst zou kunnen gebeuren. Hierbij valt veel concurrentievoordeel te behalen omdat men naar de toekomst kijkt waarbij het concurrenten een stap voor kan zijn door te anticiperen op toekomstige ontwikkelingen.

[2] Literatuurlijst: bron 1

5 Predictive Analytics

Predictive Analytics is het intelligent en doelgericht gebruiken van beschikbare data. Predictive Analytics bestaat uit een aantal technieken van statistiek en datamining die actuele en historische feiten analyseren om voorspellingen te maken van toekomstige gebeurtenissen. Marketeers die werkzaam zijn bij grote, marketinggestuurde bedrijven, schaffen steeds vaker predictive analytics-software (artikel in bijlage L). In dit hoofdstuk wordt eerst ingegaan op de theorie van datamining, vervolgens zijn er beschrijvende en voorspellende analyses uitgewerkt met korte voorbeelden.

5.1 Datamining[3]

Datamining is geschikt om op een geautomatiseerde manier patronen en relaties te ontdekken in grote hoeveelheden gegevens. Datamining kan bijna letterlijk worden vertaald als het opgraven van informatie uit databases. Zonder datamining beperkt een data-analyse zich tot een klein aantal aspecten waarvan men vermoedt dat ze een rol spelen. Met datamining wordt de scope van de analyse enorm vergroot. Door het toepassen van kunstmatige intelligentie, statistische technieken en visuele rapportagemethoden worden verborgen patronen en verbanden gevonden. Zo wordt het mogelijk om grote gegevensverzameling te analyseren en complexe verbanden te ontdekken.

Afbeelding 4 - datamining

Door datamining toe te passen op databases van bijvoorbeeld klanten, transacties of gebeurtenissen, kan snel diepgaand inzicht worden verschaft en kunnen modellen worden gebouwd die typische eigenschappen herkennen en individueel gedrag kunnen voorspellen. Deze optimale informatiepositie resulteert in betere beslissingen en meer effectiviteit op strategisch, tactisch en operationeel niveau.

5.2 Beschrijvende analyses

Beschrijvende analyses worden gebruikt om een eerste indruk van de data te krijgen en analyses uit te voeren. Hierbij valt te denken aan het aantal klanten dat een product in een bepaalde periode heeft gekocht of het gemiddeld aantal belminuten dat een klant per maand verbruikt. Hieronder staan een aantal korte voorbeelden van beschrijvende analyses.

5.2.1 Kruistabel

Een kruistabel wordt gebruikt om snel een globaal inzicht te krijgen in het verloop van en het mogelijk verband tussen twee of meer variabelen.

Voorbeeld: In de kruistabel (tabel 1) is het aantal klanten weergegeven per regio die gereageerd hebben op een mailing.

	Respons	Geen respons	Totaal
Regio 1	88	21	109
Regio 2	97	34	131
Regio 3	35	42	77
Regio 4	101	15	116
Regio 5	90	8	98
Totaal	**415**	**120**	**535**

Tabel 1 - Voorbeeld kruistabel

[3] Literatuurlijst: bron 13

Grafiek 1 - Voorbeeld

5.2.2 Grafische voorstellingen

Tabellen kunnen een belangrijk hulpmiddel zijn om een onoverzichtelijke hoeveelheid gegevens te ordenen en toegankelijk te maken voor de lezer. Een soortgelijke functie heeft de grafische voorstelling. De lezer kan vaak in één oogopslag vaststellen wat de belangrijkste conclusies zijn die volgen uit de beschikbare gegevens als het gedrag van de betrokken grootheden in een grafiek weergeven wordt.

5.2.3 Associatieregels[4]

Associatieregels zijn uitspraken over de samenhang tussen attributen, het gaat hierbij om het formuleren van waarschijnlijkheden. Op basis van gegevens van bekende, vergelijkbare consumenten wordt bijvoorbeeld de aanschaf van een item waarschijnlijk of minder waarschijnlijk geacht.

Voorbeeld: Klanten die product X kochten, kochten ook product Y en Z. Associatie modellen koppelen een bepaalde conclusie (zoals het besluit om iets te kopen) met een reeks voorwaarden.

5.2.4 Clusteranalyse[5]

Clusteranalyse is een verzamelnaam voor een groot aantal technieken en procedures. De doelstelling van clusteranalyse is om gevallen (individuen, producten, merken, onderzoekseenheden) in groepen te sorteren, zodanig dat een hoge mate van gelijkenis bestaat tussen gevallen in dezelfde groep.

Voorbeeld: Een telecom bedrijf heeft drie clusters onderscheiden: groot-, middel- en kleinverbruikers van telefoonbundels (tabel 2). Het telecom bedrijf heeft een beeld gevormd van clusters en heeft ook bepaalde responsverwachtingen per cluster. De samengestelde clusters worden vervolgens in de praktijk bewerkt met een mailing. Hierbij wordt een aanbieding gedaan op basis van telefoonverbruik met de doelstelling om de respons van klanten te verhogen.

Cluster	Aantal	Verwachte respons	Gerealiseerde respons
Groot verbruikers	10.000	20 %	14 %
Midden verbruikers	10.000	12 %	16 %
Klein verbruikers	10.000	6 %	12 %

Tabel 2 - Voorbeeld clusteranalyse

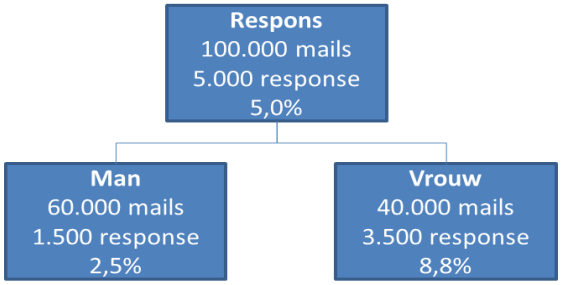
5.3 Voorspellende analyses

Een voorspellende analyse maakt een schatting van gebeurtenissen in de toekomst, gebaseerd op feiten in het verleden. Deze voorspellende technieken zijn uitermate geschikt als een organisatie bijvoorbeeld wil weten welke klanten een bepaald product in de toekomst zullen gaan afnemen of welke klanten een hoge kans hebben om over te stappen naar de concurrentie.

5.3.1 Beslisbomen

Een Beslissingsboom is een model dat een voorspelling doet op basis van een reeks beslissingen. Elke tak van de boom is een classificatievraag en de bladeren van de boom stellen delen van de gegevensverzameling voor.

Voorbeeld: In de beslisboom in afbeelding 5 is het aantal klanten dat reageert op bepaalde e-mails uitgewerkt.

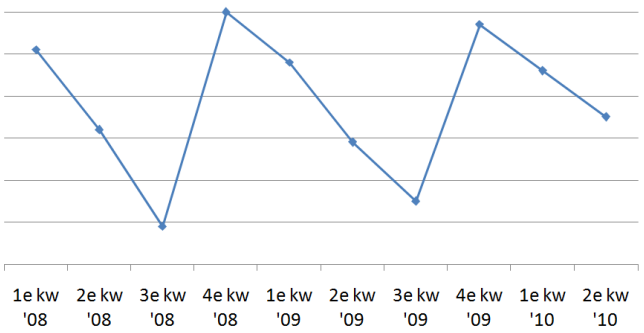


Afbeelding 5 – Voorbeeld beslisboom

[4] Literatuurlijst: bron 4
[5] Literatuurlijst: bron 6

5.3.2 Tijdreeks analyse

Een tijdreeks analyse is de analyse van gegevens die verzameld zijn in de loop van een bepaalde periode en kan voor verschillende doeleinden worden gebruikt. Zo biedt een tijdreeks analyse de mogelijkheid reeksen uit het verleden te beschrijven of de mogelijkheid tot het doen van voorspellingen in de toekomst. Een tijdreeks analyse is met name geschikt om seizoensinvloeden te achterhalen. Het verschil tussen een regressieanalyse en tijdreeks analyse is dat de uitkomst wordt gemodelleerd als een periodiek proces wat niet mogelijk is met gewone (lineaire of non-lineaire) regressie.

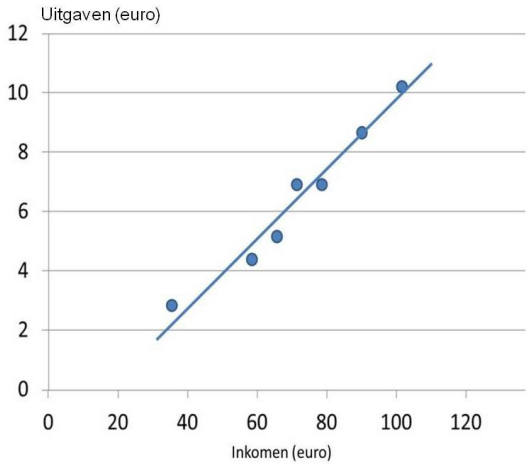


Grafiek 2 – Voorbeeld tijdreeks analyse

Voorbeeld: In grafiek 2 is een voorbeeld weergegeven van het gasverbruik per kwartaal door middel van een tijdreeks uit het verleden. Aan de hand van deze tijdreeks is het mogelijk om het gasverbruik in de toekomst per seizoen te kunnen voorspellen.

5.3.3 Regressie[6]

Regressie is een manier om de relatie tussen twee of meer variabelen tot uitdrukking te brengen. Er zijn verschillende soorten regressieanalyses beschikbaar: enkelvoudige en meervoudige regressie en lineaire en non-lineaire regressie. Het eenvoudigste model is lineaire regressie en deze heeft als doel de relatie tussen variabelen weer te geven door middel van een rechte lijn door een puntenwolk. Deze rechte wordt zo getrokken dat een zo goed mogelijke weergave van het verband tussen de variabelen ontstaat.



Grafiek 3 – Voorbeeld lineaire regressie

Voorbeeld: Er zijn gegevens verzameld over de gemiddelde dag uitgaven per persoon aan supermarkten en het gemiddeld inkomen per persoon per dag. Aan de hand van deze gegevens wordt geprobeerd een rechte lijn te trekken, die een zo goed mogelijke weergave geeft van het verband tussen de beide variabelen (grafiek 3).

5.3.4 Neurale netwerken

Een neuraal netwerk is in enige mate gebaseerd op de organisatie van het menselijke brein en de manier waarop de hersenen leren. Er zijn twee belangrijke structuurelementen in het neurale netwerk:

- Het knooppunt;
  - Te vergelijken met de zenuwcel in een menselijk brein;
- De koppeling;
  - Te vergelijken met de verbindingen tussen deze zenuwen.

Neurale Netwerken is een analyse methode die vooraf geen informatie nodig heeft over de verbanden tussen de input en output. De relaties tussen input variabelen en de output variabelen worden bepaald door de gegevens, en niet voorbepaald door een model.

In bijlage D staat beschreven welke analysetechniek je in welke situatie het beste kunt gebruiken.

[6] Literatuurlijst: bron 6

6 Analysemethodes

Door middel van analysemethoden kunnen de verzamelde gegevens worden geanalyseerd. In dit hoofdstuk worden de analysemethodes klanten- en fraudeanalyse uitgewerkt; voor de uitgebreide versie van deze analyses wordt verwezen naar bijlage E.

6.1 Klant analyse

Een klantanalyse richt zich op het analyseren, toepassen en opslaan van kennis over klanten en manieren om klanten te benaderen.
De klant is tegenwoordig koning, waardoor focussen op een succesvolle één op één relatie noodzaak is geworden. Om de maximale klantwaarde te behalen is het van belang om te weten wat de klant gaat doen en waar de klant behoefte aan heeft. Middens het toepassen van een analytische tool op deze gegevens kunnen de behoeften van de klant worden voorspeld zodat de klanten tevreden kunnen worden gehouden en de klantwaarde kan worden gemaximaliseerd.

Het voorbeeld uit tabel 3 is gebaseerd op een fictief bedrijf dat meer winstgevende resultaten in de toekomst wil bereiken door middel van de marketingcampagnes beter af te stemmen op de klant.

Segment	Score	Cover...	Frequency	Probability
⊟ **income** income > 55.267	1	643	551	85,69%
⊟ **age** age > 30 and age <= 34	1	505	111	21,98%

Tabel 3 – Beslissingslijst klantanalyse

Het voorbeeld maakt gebruik van een beslissingslijst model om de kenmerken van klanten die het meest reageren te identificeren, gebaseerd op voorgaande acties. Aan de hand van deze lijst is te zien dat klanten met een hoger inkomen dan 55.267 het beste kunnen worden benaderd want hiervan reageren 551 van de 643 wat neerkomt op 85,7%.

6.2 Fraude[7]

Frauderen is niets anders dan een vorm van diefstal. Waar diefstal traditioneel geassocieerd wordt met het ontvreemden van materiële goederen, met name liquide geld, wordt fraude veelal in verband gebracht met misdadigheid in kringen waar veel geld in omgaat. Voorbeelden hiervan zijn het stelen van informatie met het oog op chantage, of het vervreemden van geld via illegale financiële transacties.

Banken, verzekeraars en overheidsinstanties signaleren zowel een toename in het aantal als in de geraffineerdheid van frauduleuze activiteiten. Om fraude effectief te bestrijden moeten organisaties continue hun processen verbeteren om klantgedrag in verschillende accounts en systemen te monitoren. Daarbij wordt onder meer gebruik gemaakt van voorspellende modellen.

De doelstelling is om door middel van geavanceerde statistische en datamining technieken, fraude proberen te bestrijden en te voorkomen. Door deze technieken worden grote hoeveelheden data geanalyseerd om zo tijdig patronen en afwijkingen op te sporen. Hiermee worden voorspellende modellen ontwikkeld die vaststellen welke klanten of activiteiten mogelijk frauduleus zijn. De inzet van predictive analytics zorgt ervoor dat er automatisch voorspellingen en aanbevelingen worden gedaan waarop beleidsmakers tijdig kunnen anticiperen.

[7] Literatuurlijst: bron 25

7 SAP BO Predictive Workbench[8]

Gedurende het onderzoek wordt gebruik gemaakt van de tool SAP BO Predictive Workbench. SAP BO Predictive Workbench is software voor geavanceerde voorspellende analyses. In dit hoofdstuk worden de mogelijkheden van SAP BO Predictive Workbench beschreven. In de eerste paragraven wordt inhoudelijk ingegaan op het ERP pakket SAP, BusinessObjects, SAP BO Predictive Workbench en Xcelsius. Vervolgens wordt er een beschrijving gegeven van de beschikbare bronsystemen, de functies en als laatste de uitkomst van de tool.

7.1 SAP

SAP is wereldwijd marktleider in softwareoplossingen voor zowel kleine en middelgrote bedrijven als internationale ondernemingen. De oplossingen van SAP helpen organisaties hun bedrijfsvoering te stroomlijnen en aanzienlijke efficiencyverbeteringen te bereiken binnen hun toeleveringsketens en bedrijfsprocessen. SAP heeft meer dan 25 branche specifieke oplossingen voor markten variërend van high tech industrie en retail tot overheids- en financiële instellingen.

Het belangrijkste product van SAP is SAP Enterprise Resource Planning (ERP). Met een ERP-pakket kunnen in principe alle bedrijfsfuncties worden ondersteund.
SAP is een informatie- en besturingssysteem waarin bedrijfsmatige processen kunnen worden vastgelegd en beheerd. Deze processen zijn ondergebracht in modules. De gegevens in de modules worden onderling direct uitgewisseld, waardoor een volledig geïntegreerd systeem ontstaat.

7.2 BusinessObjects

SAP BusinessObjects is een SAP oplossing en bevat verschillende BI oplossingen. SAP BusinessObjects omvat een groot aantal tools en toepassingen waarmee de prestaties van een onderneming kunnen worden verbeterd.

De Business Intelligence oplossingen van SAP BusinessObjects voorzien de gebruiker van tools die ze nodig hebben om zinvolle en doordachte beslissingen te baseren op solide gegevens en analyses. Hierdoor wordt iedereen, van de topanalist tot de incidentele zakelijke gebruiker, van de benodigde informatie voorzien, zonder al te veel afhankelijkheid van IT afdelingen en ontwikkelaars.

Met deze krachtige oplossingen kunnen gebruikers in de gehele onderneming informatie vinden, bekijken, analyseren en delen met medewerkers van de organisatie.
Enkele kenmerken van de business intelligence van SAP BusinessObjects:

- Geavanceerde analysetools;
- Dashboards en Visualisatie;
- Informatie Infrastructuur;
- Query en Analyse;
- Rapporten;
- Zoeken en navigatie.

[8] Literatuurlijst: bron 10

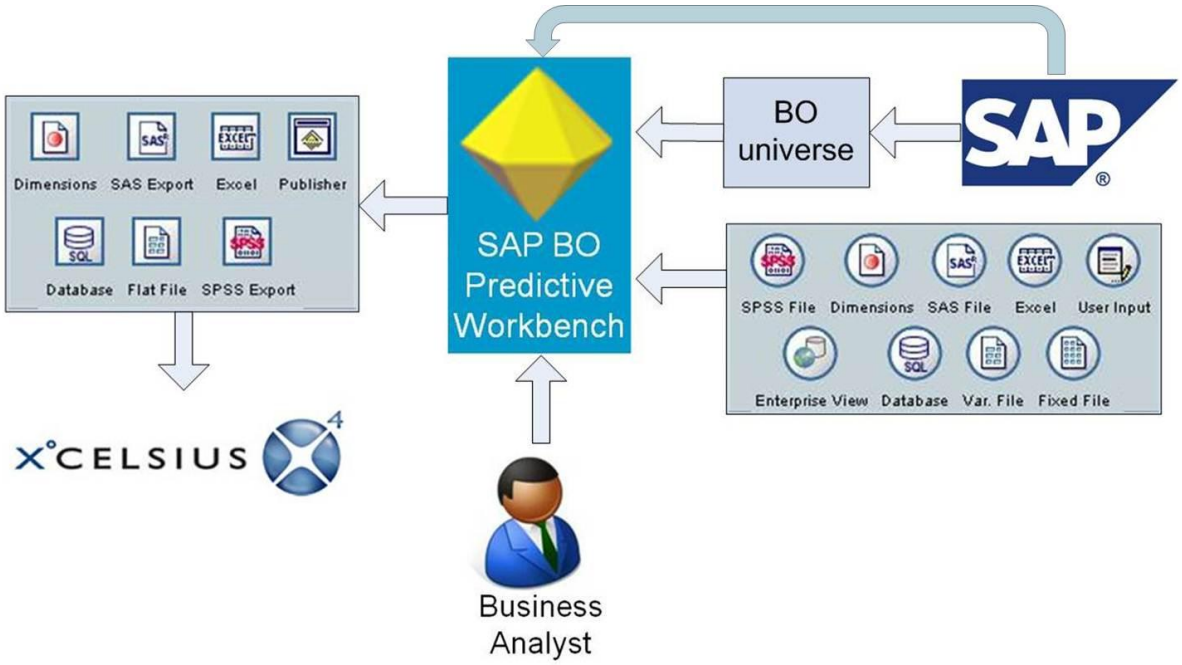
7.3 Tool

De Tool SAP BO Predictive Workbench is een onderdeel van SAP BusinessObjects en laat analyses los op de gegevens uit deze omgeving. SAP BO Predictive Workbench is een software voor geavanceerde voorspellende analyses en gebaseerd op PASW Modeler van SPSS, een van de meest krachtige en bekende tools waarmee bedrijfsanalisten data analyseren. De tool is door de eenvoudige grafische interface makkelijk aan te passen en te gebruiken maar kennis van statistiek of analytics is hierbij wel noodzakelijk om de werkwijze en de uitkomst van de analysemodellen te begrijpen.
De gemaakte analyses kunnen wel gevisualiseerd worden waardoor de uitkomst voor alle eindgebruikers op een snelle manier inzichtelijk wordt en kan worden bekeken waardoor er vervolgens snel op kan worden gereageerd.
De software kan voor deze analyses bovendien actuele gegevens uit diverse externe bronnen betrekken.

7.3.1 Beschikbare bronsystemen

In afbeelding 6 is de data flow binnen de architectuur van SAP BO Predictive Workbench weergegeven. SAP BO Predictive Workbench draait op het BusinessObjects Enterprise platform en gebruikt universes. Vanuit SAP kunnen gegevens in de BusinessObjects universe worden gezet en vervolgens ingelezen worden in de tool SAP BO Predictive Workbench. Hiernaast bestaat ook de mogelijkheid om uit verscheidende bronnen de gegevens te halen bijvoorbeeld uit een Excel bestand, SQL database, tekst bestand of direct uit SAP BW via query's.

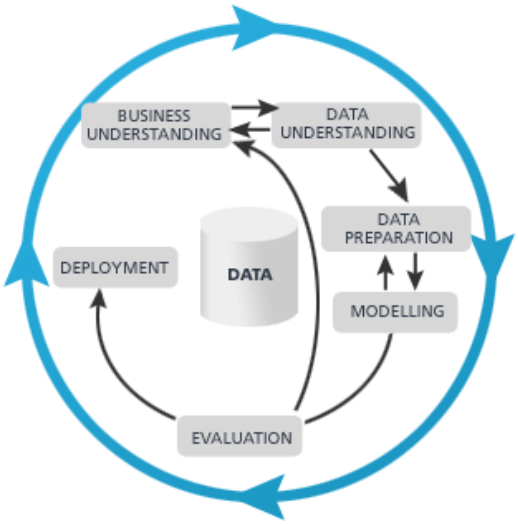
In de tool SAP BO Predictive Workbench worden deze gegevens verwerkt tot voorspellende analyses die de business analist nodig heeft om gegevens te analyseren. De uitkomst van deze voorspellende analyses kan worden weergeven in bijvoorbeeld een tabel, Excel bestand of rapport. De uitkomst van de analyses kan vervolgens in Xcelsius worden gezet waarmee de resultaten beter gevisualiseerd kunnen worden. Xcelsius bevat een intuïtieve interface met opvallende visualisatiecomponenten en gegevens connectoren waarmee complexe bedrijfsgegevens in gemakkelijk te begrijpen dashboards kunnen worden gezet en getoond.



Afbeelding 6 - Dataflow SAP BO Predictive Workbench

7.3.2 Methode[9]

De methode die wordt gevolgd bij de tool SAP BO Predictive Workbench voor het maken van analyses is de methodologie van CRISP-DM.
CRISP-DM staat voor CRoss Industry Standard Process for Data Mining. Dit is een procesmodel dat ervoor zorgt dat grote datamining projecten sneller, goedkoper, betrouwbaarder en handelbaar blijven.



Afbeelding 7 - Het CRISP-DM model

Stappen procesmodel

CRISP-DM is een iteratief proces waarbij stappen dus kunnen worden herhaald, wat vaak gebeurt door bijvoorbeeld meerdere technieken op de dataset los te laten, of er kan worden teruggegaan naar een eerdere stap. De volgende stappen worden onderscheiden:

Opstartfase

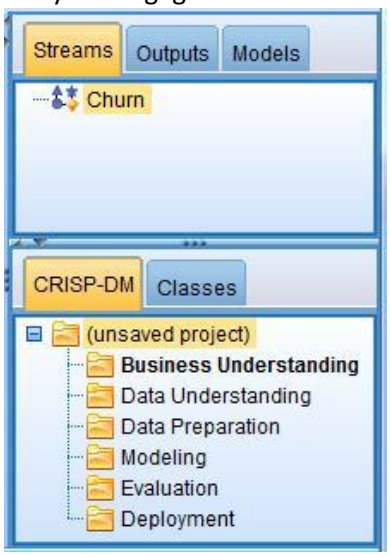
Als eerste wordt een business requirements document opgesteld waarin specifiek een probleem wordt omschreven dat moet worden opgelost.

Gegevens oriëntatie

Deze fase bestaat uit het selecteren van de data om het probleem op te kunnen lossen, het verkennen en het beschrijven van de data.

Gegevenspreparatie

Nadat de juiste gegevens zijn geselecteerd moeten ze worden geprepareerd voor de werkelijke analyse. De gegevens moeten worden opgeschoond omdat vaak waarden ontbreken en er fouten in zitten. Tevens moeten de gegevens in de juiste vorm in één tabel worden gezet en moeten vaak transformaties (omzettingen) op de tabel plaatsvinden. Deze zijn nodig om extra informatie toe te voegen zoals ratio's en sommaties (bijvoorbeeld per klant).



Afbeelding 8 - CRISP-DM

Modelleringsfase

In deze stap wordt de analyse op de gegevens losgelaten. Vaak wordt in de tool een techniek geselecteerd, waarna een aantal parameters van de gekozen techniek moeten worden ingesteld. Vervolgens kan het model worden uitgevoerd. Er wordt op deelverzamelingen van de data getest om te kijken wat goede modellen zouden kunnen zijn (bijlage F).

Evaluatiefase

In deze fase moet worden afgevraagd of het vereiste doel van de analyse behaald is en of herhalingen van het uitgevoerde proces nodig zijn.

Actiefase

Voor de gevonden resultaten moet een strategie worden bepaald. Er moet worden gezorgd dat de informatie bij de juiste personen terechtkomt en dat acties worden genomen om het probleem op te lossen.

[9] Literatuurlijst: bron 8

7.3.3 Mogelijkheden

Met SAP BO Predictive Workbench is het onder andere mogelijk om te voorspellen en te plannen. De krachtige oplossing van Predictive Workbench kan worden gebruikt als 'glazen bol' om de organisatie alert te houden en de concurrentie één stap voor te blijven. De kracht van voorspellende analyses kan veel industrie specifieke problemen oplossen zoals:

- Het voorspellen van garantieclaims voor producten;
- Het vinden van de juiste prijsstelling voor klantsegmenten;
- Het voorspellen van kosten.

Door middel van de tool kan ook zicht worden verkregen op het te verwachten gedrag van klanten. Kennis die kan worden gebruikt om met minder middelen meer te bereiken.
Deze kennis kan de onderneming helpen om slimmer te werken, bijvoorbeeld:

- Klanten die niet zullen reageren ook niet benaderen;
- Kleinere maar zeer doelgerichte campagnes behouden met wellicht en even hoge response;
- Klanten naar klantwaarde te segmenteren;
- Het prijsbeleid optimaliseren.

Daarnaast is het ook mogelijk om trends en patronen te ontdekken in alle data die zijn opgeslagen in diverse interne en externe databronnen. Hiermee kunnen nieuwe en oude feiten weer worden ontdekt waardoor betere beslissingen kunnen worden genomen en organisatiedoelen worden bereikt.

7.3.4 Resultaat

Met SAP BO Predictive Workbench worden voorspellende analyses gebruikt met innovatieve en bewezen BI mogelijkheden. Resultaten van SAP BO Predictive Workbench kunnen worden opgeroepen vanuit SAP BusinessObjects rapport of – Xcelsius dashboard.
Gebruikers zien bijvoorbeeld de lijst met winstgevende klanten binnen SAP BusinessObjects rapporten. Ook kunnen de KPI's (Key Performance Indicators) voor klantwinstgevendheid worden gecontroleerd.

Voordelen die onder andere kunnen ontstaan met de tool SAP BO Predictive Workbench zijn:

- Betere en tijdige fraudedetectie;
- Eenvoudige en snellere klantenservice processen;
- Hogere klanttevredenheid;
- Hogere conversie van campagnes of folders;
- Betere segmentatie van klanten;
- Besparing op direct marketing;
- Kostenefficiënt, gericht en effectief benaderen van klanten.

Xcelsius

Met Xcelsius kunnen de data van een organisatie op een eenvoudige manier worden veranderd in aantrekkelijke presentaties met dynamische grafieken, grafische voorstellingen en interactieve visuele analyses. Hierbij kunnen schuifregelaars en wijzers worden gebruikt om onderliggende veronderstellingen te veranderen. De resultaten worden onmiddellijk en dynamisch op het scherm weergegeven.

8 Branches[10]

Vele ondernemingen en overheidsinstellingen zouden door middel van SAP BO Predictive Workbench inzichten naar boven kunnen halen, die verborgen liggen in de gegevens. Hierdoor kunnen de bedrijfsresultaten verbeterd worden en gestelde doelen uiteindelijk beter worden bereikt.

SAP BO Predictive Workbench kan de volgende branches helpen:

Bedrijven voor bancaire en financiële diensten	Om marketingcampagnes doelmatiger te maken, om op een meer betrouwbare wijze kredietrisico's te beoordelen en om mogelijk frauduleuze activiteiten doelmatiger op te sporen.
Verzekerings-maatschappijen	Om marketinginspanningen nog beter af te stemmen en de verwerking van claims te stroomlijnen, en om frauduleuze claims op een meer betrouwbare wijze te onderkennen.
Telecommunicatie-bedrijven	Om hechtere banden met klanten te ontwikkelen om zo een betere loyaliteit op te bouwen en het aantal vertrekkende klanten ('verloop') te reduceren.
Detailhandelaars	Om de planning van het assortiment te verbeteren en de inspanningen voor marketing en op het gebied van de loyaliteit van klanten beter af te stemmen.
Nutsbedrijven en energieleveranciers	Om klanten een meer persoonlijke service te bieden. Analyses spelen ook een rol bij preventief onderhoud, hetgeen uitmondt in een grotere betrouwbaarheid tegen lagere kosten.
Overheids-instellingen	Bij het beheer van allerlei functies die variëren van selecties voor belastingcontroles en werving voor legeronderdelen tot proactieve politiediensten en openbare veiligheid.
Organisaties in de gezondheidszorg	Om voorspellende intelligentie te gebruiken voor een proactief beheer van resources en een precieze afstemming van hun werkpraktijken voor een betere patiëntenzorg.
Hogescholen en universiteiten	Om de gehele cyclus die studenten doorlopen, doelmatiger te beheren, van het werven van de juiste mix aan studenten tot het aanbieden aan hen van een selectie aan programma's en soorten hulpverlening om ze ingeschreven te houden, en een beter geslaagd beheer van programma's voor verdere ontwikkeling van alumni.

Met de SAP BusinessObjects oplossing Predictive Workbench kan een bedrijf zich ontwikkelen tot een 'Predictive Enterprise'. Hiermee wordt bedoeld dat een organisatie beslissingen stuurt en optimaliseert door het integreren van voorspellende technologie met als doel de winstgevendheid te maximaliseren.

[10] Literatuurlijst: bron 16

9 Werkzaamheden onderzoek

In dit hoofdstuk zijn de werkzaamheden uitgeschreven die tijdens de afstudeeropdracht zijn uitgevoerd. De volgende paragraven gaan eerst over het ontdekken van de tool en het maken van de demo waarna er wordt ingegaan op de realisatie van de koppelingen met BusinessObjects en Xcelsius.

9.1 Onderzoek tool

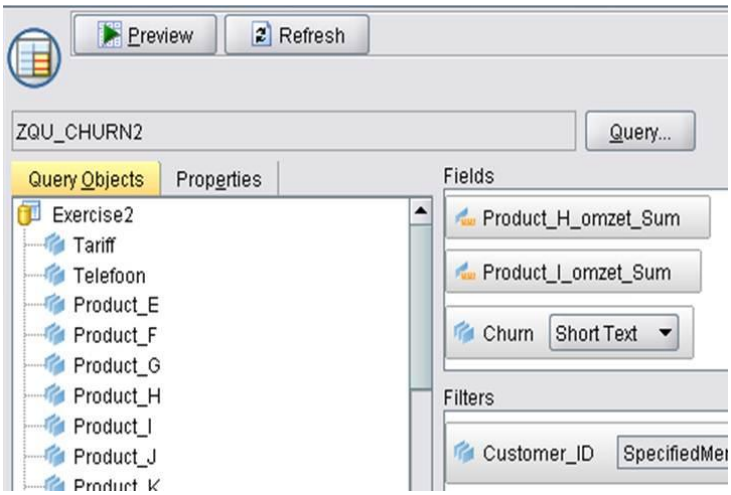
Na de theorie van de tool uitgewerkt te hebben is de tool SAP BO Predictive Workbench geïnstalleerd. Dit was een stand alone versie die door medewerkers van Accenture erop is gezet en het draaide meteen. Het enige nadeel was dat er nog geen licentierechten voor de tool waren, waardoor een tijdelijk licentie werd gebruikt die slechts voor een periode van 2 weken werkten. Na deze periode werd steeds een virtuele machine aangemaakt waardoor de tool wederom 2 weken te gebruiken was. Ondanks deze terugkerende handeling heeft de tool heel de afstudeerperiode goed gewerkt.

Nadat de tool beschikbaar was, is gekeken naar de mogelijkheden van de tool. Er was geen handleiding van de tool beschikbaar maar wel een uitgebreide help mogelijkheid waar veel op te vinden was. De tool zelf bevatte een aantal verschillende demo's met uitleg erbij, waardoor de tool intuïtief kon worden ontdekt. Door bij elke demo uit te zoeken wat per stap precies gebeurde en werd berekend, werd de werkwijze van de tool al snel duidelijk.

Nadat een aantal demo's volledig duidelijk waren, kon gestart worden met een eigen demo. Doordat de demo gegevens beschikbaar waren van een telecom bedrijf is er voor gekozen om dit onderwerp ook uit te werken in de business case. In de demo zijn eerst alle gegevens op een logische manier bewerkt en samengevoegd. Na de bewerking van de gegevens konden analysemodellen worden ingezet om verschillende analyses te maken. De tool gaf steeds aan dat de uitkomsten van de modellen kwalitatief erg goed waren, waardoor de analyse ook een bepaalde zekerheid bood. Dit kwam waarschijnlijk doordat de demogegevens kwalitatief goed waren en de juiste bewerking was toegepast. De demo is uiteindelijk volledig uitgewerkt waarbij verschillende modellen op de data zijn uitgevoerd en vervolgens uitgewerkt.

9.2 Koppelingen

In de demo werd gebruik gemaakt van tekst, Excel en Acces bestanden die als brongegevens werden in gelezen in de tool. Het doel van SAP BO Predictive Workbench is om een rechtstreekse koppeling te maken met BusinessObjects. Om dit te realiseren moesten eerst gegevens in SAP BW (paragraaf 9.2.1) worden ingelezen waar vervolgens in BusinessObjects een universe van kon worden gemaakt.

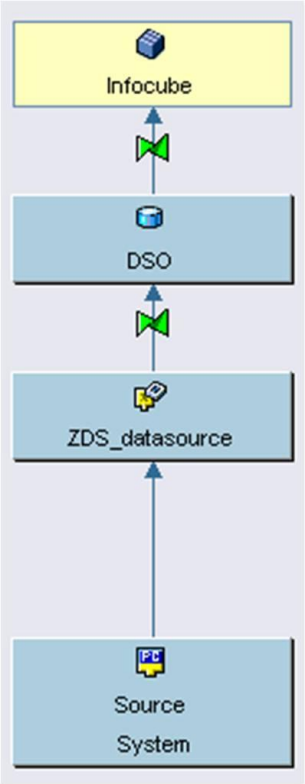


Afbeelding 9 - Koppeling met universe vanuit SAP BO

9.2.1 SAP Business Warehouse[11]

Om de gegevens van de telecom demo in te laden in SAP is een datawarehouse van deze gegevens in SAP BW gebouwd. Vanwege de SAP Business Warehouse lessen tijdens de opleiding en het volgen van een bootcamp bij Accenture over het onderwerp kon dit onderdeel als extra worden uitgevoerd.

Een datawarehouse is meestal een gigantische hoeveelheid aan elektronische data die bedoeld is om beslissingen te nemen. Deze gegevens moeten grondig geanalyseerd worden aangezien ze gebruikt worden als basis voor belangrijke beslissingen.
Om de gegevens in SAP BW te kunnen zetten moet eerst een datasource van de telecom data worden aangemaakt. Het geselecteerde (in dit geval csv) bestand wordt ingelezen. Na het checken en activeren van de DataSource worden alle ongewijzigde data vanuit de datasource ingelezen en opgeslagen in de PSA (Persistent Staging Area). PSA wordt gebruikt als een tijdelijke opslag van het bron systeem. De gegevens staan dan centraal in de PSA op de SAP BW server. Door die tijdelijke opslag is er nog de mogelijkheid om de data te controleren voordat de gegevens in het datawarehouse van SAP BW gezet worden.
Vervolgens wordt een DataStore Object aangemaakt waarin de gegevens worden gekopieerd vanuit de Datasource en gecontroleerd op data kwaliteit.
Als laatste wordt een InfoCube gecreëerd waarmee in feite de bouw van het sterschema plaats vindt. Een sterschema zorgt ervoor dat gegevens uit verschillende tabellen gecombineerd worden in slechts enkele tabellen. Hierdoor hoeven gebruikers slechts enkele tabellen aan te spreken om analyses te plegen, zodat die veel sneller gemaakt kunnen worden. Van de gegevens in SAP BW kan een query worden gemaakt waarin de benodigde gegevens worden gezet voor een bepaalde analyse. Deze query kan vervolgens worden geopend in BusinessObjects waarna hiervan een universe kan worden gemaakt.



Afbeelding 10 - Dataflow

Het is tijdens deze afstudeeropdracht gelukt om zowel rechtstreeks vanuit de SAP BW (query) als vanuit de universe in BusinessObjects een koppeling te maken vanuit SAP BO Predictive Workbench. Door middel van deze koppeling werkte de demo in de tool op dezelfde wijze alleen dan met gegevens rechtstreeks vanuit BusinessObjects of SAP BW.

9.2.2 Onderzoek Xcelsius

Na een koppeling van de telecom data met SAP BW en daarna met Business Objects te hebben gerealiseerd was een volgende extra opdracht om een koppeling met Xcelsius te maken. Doordat de voorspellingen in de demo voornamelijk gingen over de kans op wel of geen churn was het niet zo interessant om deze uitkomst in Xcelsius te zetten. Om toch de mogelijkheden van Xcelsius weer te geven zijn er analyses gemaakt van de gegevens van de demo. Door in het begin te inventariseren welke rapportages gemaakt moesten worden, konden de gegevens op de juiste manier worden gerangschikt in Xcelsius. Dit maakte het opstellen van de rapportages wat gemakkelijker en zorgde voor een mooi resultaat. In bijlage K staat een voorbeeld weergegeven met de volledige uitwerking.

[11] Literatuurlijst: bron 24

10 Business case

De business case gaat over de Churn (weglopen van klanten) in de organisatie en op welke manier de churn het beste kan worden voorspeld door middel van SAP BO Predictive Workbench. In paragraaf 1 wordt de probleemdefinitie uiteen gezet en hierbij wordt de telecom branche als voorbeeld gebruikt. In paragraaf 2 worden de mogelijkheden beschreven waarmee het probleem opgelost kan worden. In paragraaf 3 wordt een investeringsbeoordeling gegeven aan de hand van een investeringsberekening met alle kosten. In paragraaf 4 zijn de belangrijkste risico's terug te vinden die bij dit project komen kijken.

10.1 Probleemdefinitie

De manier waarop een organisatie met zijn klanten omgaat, wordt steeds belangrijker. Klanten worden veeleisender en willen zelf steeds meer invloed uitoefenen op financiële diensten en producten. In combinatie met een groeiende diversiteit aan communicatie- en verkoopkanalen is het voor bedrijven tegenwoordig moeilijk om klanten te werven en te behouden.
Het gevolg hiervan is dat klanten makkelijker overstappen van de ene naar de andere leverancier, waardoor het van groot belang is om te weten op welke punten ze tevreden zijn of juist niet en welke klanten dreigen te vertrekken.

Het voorspellen van churn is erg belangrijk voor een organisatie. Het is namelijk goedkoper om een bestaande klant te behouden dan om een nieuwe klant te werven. Dat maakt het tegengaan van klantverlies voor veel organisaties zinvol.

10.1.1 Voorbeeld Telecom branche

Als voorbeeld is in deze business case de telecom branche uitgewerkt. Deze branche is op dit moment volop in beweging. De markt kent vele miljoenen afnemers, die steeds hogere eisen stellen aan het product en de dienstverlening.
De telecom markt is een snel veranderende en competitieve markt. Het voornaamste doel van een telecom bedrijf is verhoging van de gemiddelde omzet per gebruiker en vergroting van het marktaandeel. Hierbij is een keuze van de juiste type telefoons, afzetkanalen en subsidieniveaus van belang. Uiteraard moet dit in de juiste verhouding staan met de opbrengsten per klant. Klanten weten steeds beter waar ze de beste aanbieding kunnen krijgen waardoor het voor een bedrijf essentieel is dat een investering die wordt gemaakt bij een klant ook wordt terugverdiend.

Klanten winnen in de markt voor mobiele telefonie is lastig en kostbaar waardoor het een grote uitdaging is om de reeds geworven klanten te kunnen behouden. Binnen deze branche heerst namelijk een enorm verloop van de ene naar de andere aanbieder, de zogenoemde churn.

Churn is een veel gebruikte term voor het afhaken van klanten in een organisatie en is een enorm verlies. Doordat uit weggelopen klanten geen omzet meer kan worden behaald, dienen deze dus snel te worden vervangen door nieuwe klanten om de omzet op peil te houden.

10.2 Oplossingen

In de vorige paragraaf is het cruciale belang beschreven om de retentie (behoud) van klanten te verhogen. In deze paragraaf worden de mogelijkheden van de tool SAP BO Predictive Workbench uiteen gezet waarmee het onder andere mogelijk is om voorspellingen te doen. Aan de hand van deze tool kunnen analyses worden gemaakt waardoor er meer inzicht bestaat in het klantgedrag zodat de mogelijke churn beter kan worden voorspeld. Wanneer een betere voorspelling kan worden gemaakt van de klanten die mogelijk gaan vertrekken, kan hierop proactief worden gereageerd door middel van bijvoorbeeld een gerichte campagne. Met het benaderen van deze specifieke klanten kan mogelijk een verhoging van de retentie worden gerealiseerd.

10.2.1 Voorspellende analyse

In deze paragraaf worden voorbeelden gegeven van de mogelijkheden van de tool SAP BO Predictive Workbench waarbij de churn beter kan worden voorspeld door middel van bepaalde voorspellende analysemodellen.

Modelleren

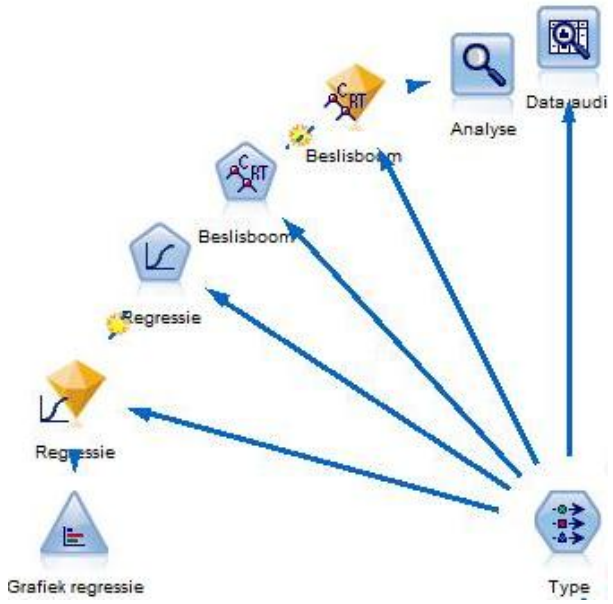
In Predictive Workbench worden een verscheidenheid van modellering methodes aangeboden waarmee nieuwe informatie uit de gegevens kan worden gehaald en voorspellende modellen kunnen worden ontwikkeld. Elke methode heeft bepaalde sterke punten en is het meest geschikt voor bepaalde type problemen.

Voor het bouwen van een model zijn minimaal drie elementen nodig:

- Een bron knooppunt die de gegevens van een externe bron inleest;
- Een type knooppunt dat veldeigenschappen zoals data type specificeert;
- Een modellering knooppunt dat een model genereert als de stroom van gegevens wordt uitgevoerd.

Daarnaast zijn ook tabel en analyse knooppunten die gebruikt worden om een overzicht van de resultaten te weergeven nadat het model is gemaakt.

Hieronder wordt een voorbeeld gegeven van het voorspellen van churn met de analysemodellen: 'Beslisboom' en 'Logistische regressie'. De volledige uitwerking van analysemodellen met betrekking tot churn staat weergegeven in de demo in bijlage J.

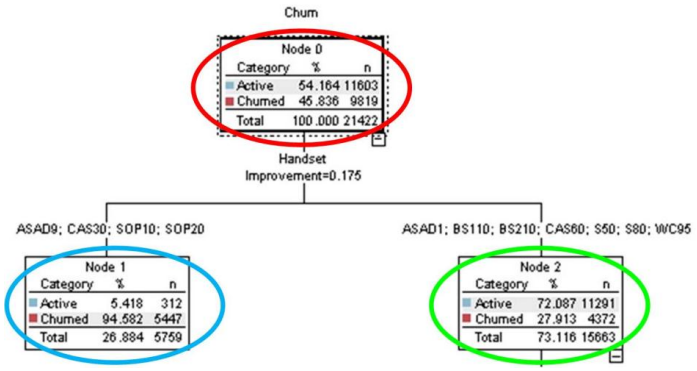


Afbeelding 11 - Analysemodellen

Beslisboom

De beslisboom (paragraaf 5.3.1) is opgesteld aan de hand van de inputwaarden tarief, geslacht, leeftijd en telefoon. Door middel van deze inputwaarden en de churn als doel voor dit analysemodel op te geven, bouwt de tool SAP BO Predictive Workbench zelf de beslisboom op. In afbeelding 12 staat het resultaat weergeven waarbij eerst node 0 (rode cirkel) zichtbaar is. Deze node heeft nog betrekking op alle klanten en de kans op churn is dan 46%.

Vervolgens wordt een opsplitsing gemaakt aan de hand van de verschillende telefoons. Wanneer klanten een telefoon hebben die bij node 1 (blauwe cirkel) is geselecteerd, is de kans op churn 95%. Terwijl deze kans bij node 2 met andere telefoons slechts 28% is. Omdat de kans op churn zo hoog is bij node 1 wordt de beslisboom hieronder niet verder opgesplitst. Dit in tegenstelling tot node 2 waar aan de hand van andere input velden de kans op churn nog verder wordt uitgediept. Hierdoor kan aan de hand van eigenschappen van klanten worden gekeken hoe groot de kans op churn is.



Afbeelding 12 - Beslisboom churn

Wanneer een analyse van de beslisboom wordt gemaakt, geeft SAP BO Predictive Workbench het resultaat van de juistheid van het model weer. De analyse in dit voorbeeld laat zien dat 87% correct is voorspeld. Aan de hand van dit percentage is vast te stellen dat de betrouwbaarheid van dit model hoog is.

Logistische Regressie

Logistische regressie wordt specifiek gebruikt om binomiale uitkomsten te modeleren. Het is gelijk aan de non-lineaire regressie waarbij een functie wordt gezocht die het best bij de data in de input dataset past. Deze functie is non-lineair ten opzichte van input waardes (voorspellende factoren) en geeft als uitkomst een waarde tussen 0 en 1 terug. Deze waarde is dus de kans dat de gebeurtenis die gemodelleerd wordt, zich voordoet (in ons voorbeeld churn). De uitkomst van het gegenereerde model is een classificatie tabel waar op basis van de voorspelde kansen wordt bepaald welke percentage van input gevallen correct is voorspeld.

Classification Table(a)

	Observed		Predicted		
			Churn		Percentage Correct
			Active	Churned	
Step 1	Churn	Active	15139	1370	91,7
		Churned	3138	10943	77,7
	Overall Percentage				85,3

a. The cut value is ,500

Tabel 4- Logistische regressie - Classificatie tabel

Logistische regressie werkt door middel van het opbouwen van een set van vergelijkingen. Zodra het model is gegenereerd, kan het gebuikt worden om waarschijnlijkheden te voorspellen voor toekomstige doeleinden.

In tabel 4 is te zien dat wanneer het model logistische regressie wordt gebruikt om de churn te voorspellen, 77,7% van de klanten die weg zullen lopen goed voorspeld wordt. Van de klanten die blijven is dit percentage 91,7%.

10.3 Investeringsbeoordeling

In dit hoofdstuk wordt er een inschatting gemaakt van de kosten en de (financiële en niet financiële) baten bij de implementatie van de tool Predictive Workbench naar aanleiding van de implementatie (bijlage G), voordelen en nadelen (bijlage H), kosten en de investeringsberekening.

10.3.1 Kosten[12]

In deze paragraaf is een voorbeeld uitgewerkt waarin de klantwaarde wordt berekend aan de hand van de omzet en de kosten die een bedrijf maakt. De volledige berekening staat weergegeven in bijlage I.

Jaar	**1**	**2**	**3**	**4**	**5**
Totaal aantal klanten	*21.422*	*19.495*	*18.409*	*17.823*	*17.537*
Gemiddelde besteding klant	€ 360	€ 360	€ 360	€ 360	€ 360
Omzet	**€ 7.711.920**	**€ 7.018.344**	**€ 6.627.312**	**€ 6.416.224**	**€ 6.313.355**
Variabele kosten	€ 1.542.384	€ 1.403.669	€ 1.325.462	€ 1.283.245	€ 1.262.671
Acquisitiekosten	€ 535.550	€ 112.500	€ 112.500	€ 112.500	€ 112.500
Retentiekosten groep 1	€ 43.219	€ 39.332	€ 37.141	€ 35.958	€ 35.381
Retentiekosten groep 2	€ 3.535	€ 3.217	€ 3.038	€ 2.941	€ 2.894
Retentiekosten groep 3	€ 21.368	€ 19.447	€ 18.363	€ 17.778	€ 17.493
Vaste kosten	€ 200.000	€ 202.000	€ 204.020	€ 206.060	€ 208.121
Winst	**€ 5.365.864**	**€ 5.238.180**	**€ 4.926.788**	**€ 4.757.742**	**€ 4.674.296**

Tabel 5 - Klantwaarde

Toelichting kosten

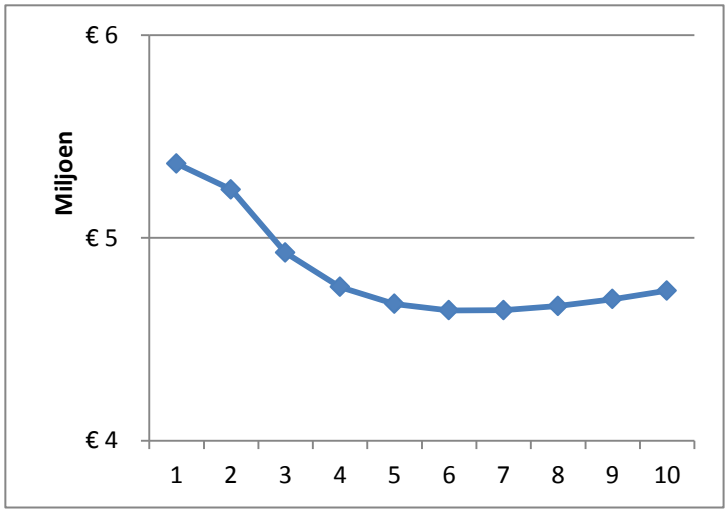
Het totaal aantal nieuw geworven klanten in jaar 1 komt uit op 21.422. De berekende retentieratio's worden losgelaten op dit aantal om te berekenen hoeveel klanten er nog over zijn in jaar 2 t/m 5.

De gemiddelde besteding per klant is €30,- per maand oftewel €360,- per jaar. In dit geval blijft de gemiddelde besteding door de jaren heen gelijk.

De omzet per jaar kan berekend worden door het totaal aantal klanten in dat jaar te vermenigvuldigen met €360,-.

De variabele kosten (kosten die veranderen door een toename of afname in de productieomvang) worden berekend door de omzet te vermenigvuldigen met 20%.

De acquisitiekosten zijn per nieuwe klant €25,-. Daarnaast zijn er nog retentiekosten en vaste kosten waarop wordt ingegaan in de volgende paragraaf.



Grafiek 4 - Winst per jaar

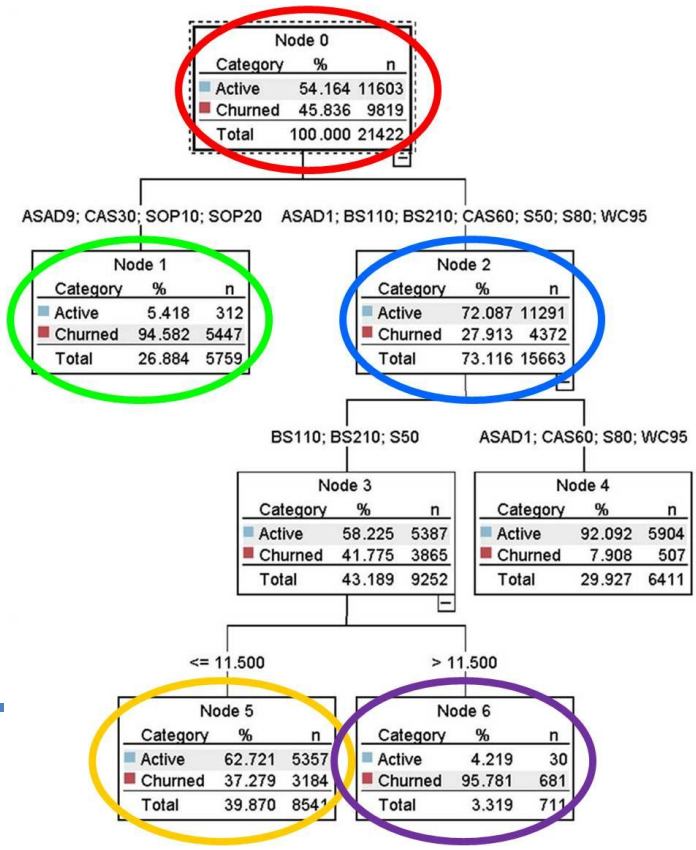
De winst per jaar wordt vervolgens berekend door de kosten van de omzet af te trekken (grafiek 4).

[12] Literatuurlijst: bron 3

Retentiekosten
De retentiekosten zijn opgedeeld in 3 categorieën. Deze categorieën zijn ontstaan naar aanleiding van een beslisboom (afbeelding 13) waar aan de hand van bepaalde inputwaarden in het model de churn wordt voorspeld. Door voor elke categorie een aparte retentie campagne te bedenken die gericht is op die groep mogelijke churn kunnen de retentiekosten worden afgestemd op die klantengroep. Hierdoor kan bijvoorbeeld de volgende verdeling van retentiekosten gemaakt worden (tabel 6).

Node	Aantal klanten	Percentage churn	Kosten per klant
1	5.759	95%	€ 7,50
5	8.541	37%	€ 2,50
6	3.319	96%	€ 5,00

Tabel 6 - Verdeling retentiekosten



Afbeelding 13 - Beslisboom kans op churn

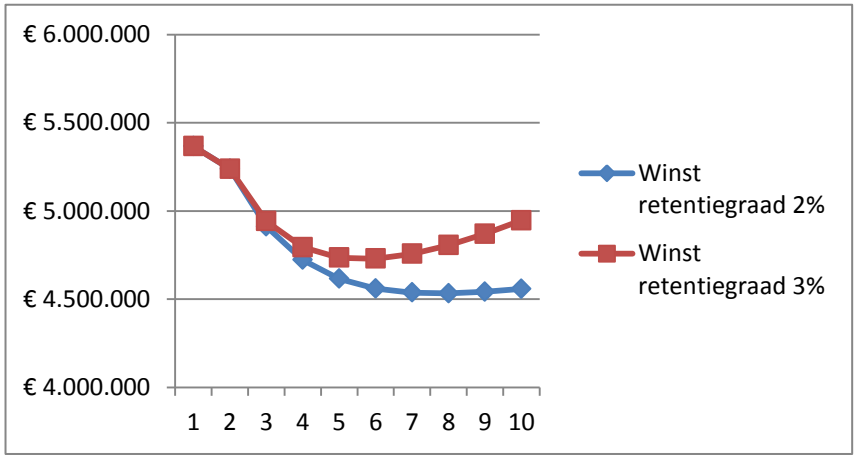
Vaste kosten
De vaste kosten bestaan uit het totaalbedrag per jaar dat ongevoelig is voor veranderingen in de omvang van de bedrijfsactiviteiten. Hieronder worden onder andere de vaste kosten van de totale implementatie van de tool SAP BO Predictive Workbench opgenomen (tabel 7).

Beschrijving	Berekening		Kosten
Cursus	€ 550,- * 5 dagen = € 2.750,- dit bedrag * 2		€ 5.500,-
Licentie tool	Bedrag gebaseerd op 2 medewerkers van het bedrijf per jaar		€ 37.500,-
Onderhoud	15% * kosten licentie per jaar *		€ 5.625,-
Totaal		**Eenmalig**	**€ 5.500,-**
		Jaarlijks	**€ 43.125,-**

Tabel 7 - Kosten tool

10.3.2 Investeringsberekening

In grafiek 5 is het verschil in winst na 10 jaar zichtbaar wanneer de retentiegraad met 1 procent wordt verhoogd. In deze grafiek wordt het belang van retentie duidelijk. Door middel van de tool SAP BO Predictive Workbench wordt de groep inzichtelijk die je kunt benaderen met een aantrekkelijke aanbieding in de hoop dat klanten langer blijven.



Grafiek 5 - Winst verschillende retentiegraden

De berekening van de investering wordt gemaakt aan de hand van de contante waarde van de onderneming. Hieronder wordt in dit voorbeeld de totale winst verstaan van een bedrijf over een tijdsperiode van 5 jaar (waarde van een bedrijf) rekening houdend met een disconteringvoet (rendementspercentage). Deze wordt meegenomen in de berekening omdat geld dat in de toekomst verdiend wordt op dit moment minder waard is dan geld dat een onderneming op dit moment in bezit heeft.
Deze berekening is uitgewerkt met 2 verschillende retentiegraden (2% en 3%) waardoor het verschil zichtbaar wordt.

Met de verhoging van de retentiegraad gaat de cumulatieve contante waarde van de onderneming van €20.753.538,- naar €20.917.971,- wat een verschil is van €164.433,-. Dit verschil wordt verkregen door de inzichten die de tool SAP BO Predictie Workbench geeft wanneer het wordt toegepast op het probleem van churn in het bedrijf.

In deze business case is er nog maar één mogelijkheid van de tool uitgewerkt, er zijn veel meer mogelijkheden om ook andere problemen op te lossen. Doordat er alleen beschikking was over demogegevens van een telecom bedrijf, is ervoor gekozen om alleen dit onderwerp gedetailleerd uit te werken.

10.4 Belangrijkste risico's

Een risico bij het implementeren van de tool SAP BO Predictive Workbench is dat de processen binnen het bedrijf niet duidelijk en hierdoor niet goed gedefinieerd zijn. Als gebruikers geen zicht hebben op de bedrijfsprocessen is er geen juist overzicht waardoor veranderingen en gestelde doelen binnen het bedrijf moeilijker uitvoerbaar worden. Het is hierdoor ook lastig om te bepalen wat er binnen het bedrijf met de tool bereikt wil worden.
Het is in eerste instantie belangrijk om de organisatie goed te kennen en precies te weten wat de doelen zijn binnen de organisatie. Door dit duidelijk uitgewerkt te hebben kan er worden gekeken of de tool toegevoegde waarde biedt voor de organisatie en op welke manier deze vervolgens kan worden ingezet.

Het grootste risico van het implementeren van de tool is dat de data niet juist, niet up-to-date of niet goed opgeschoond is waardoor de uitkomsten van de analyses niet kloppen met de werkelijkheid. Het is van belang dat de datakwaliteit hoog is en dat de data voordat deze in gebruik wordt genomen helemaal opgeschoond wordt. Deze opschoning kan worden uitgevoerd door de gegevens te uniformeren, standaardiseren, aanvullen, ontdubbelen, splitsen en spreiden.
Door de data op deze manier op te schonen en actuele waarden te gebruiken wordt de datakwaliteit verhoogd wat leidt tot betrouwbaardere modellen.

In het voorbeeld van deze business case heeft dit risico zich niet voorgedaan omdat de demo data van goede kwaliteit was en geen lege of dubbele waarden bevatte. De tool zelf gaf ook aan dat de uitkomst van de analyses erg betrouwbaar was met een goede kwaliteit.

Dat de tool gebruiksvriendelijk is, is geen garantie dat iedereen er zonder achtergrond mee kan werken. Inhoudelijke kennis van de tool en analytics is noodzakelijk om deze op een juiste manier te gebruiken.

11 Conclusie(s) en aanbeveling(en)

De afstudeeropdracht is conform het PID uitgevoerd. Het is zelfs mogelijk geweest om ook 'extra' opdrachten (koppeling en visualisatie) die in het PID stonden vermeld uit te voeren. De tool SAP BO Predictive Workbench is draaiend gekregen waardoor de mogelijkheden van de tool verkend konden worden. Na deze mogelijkheden uitgezocht en ontdekt te hebben, kon er een demo worden opgesteld. Voor deze demo werd er gebruik gemaakt van gegevens uit een Excel-, tekst- en Acces bestand. Nadat deze demo helemaal was afgerond en werkte, is het uiteindelijk ook gelukt om deze gegevens door middel van een koppeling met SAP BW (query) en BusinessObjects in SAP BO Predictive Workbench te krijgen. Van de demo gegevens uit BusinessObjects is uiteindelijk nog een voorbeeld gegeven van de visualisatie mogelijkheden van Xcelsius.

De tool SAP BO Predictive Workbench maakt het voor gebruikers mogelijk om grote hoeveelheden beschikbare data snel en betrekkelijk eenvoudig om te zetten in analyses.
De tool is door de eenvoudige grafische interface gemakkelijk te gebruiken maar kennis van statistiek en analytics is hierbij wel noodzakelijk. De grafische interface leidt gebruikers snel door het gehele data mining-proces, van datavoorbereiding tot aan modelontwikkeling.
De gemaakte analyses kunnen gevisualiseerd worden waardoor het resultaat voor alle eindgebruikers op een snelle manier inzichtelijk wordt. Wanneer bekend is welke factoren invloed hebben op de resultaten, kunnen kansen worden benut en risico's worden beperkt. Aan de hand van deze resultaten kan vervolgens proactief worden gereageerd zodat bijvoorbeeld marketingcampagnes kunnen worden geoptimaliseerd. Hierdoor kunnen de productiviteit, de flexibiliteit en de prestaties van een bedrijf worden verbeterd.

Wanneer een organisatie SAP BO Predictive Workbench in wil zetten, worden diverse eisen gesteld aan het aan de kwaliteit van de data. Hierbij is het belangrijk dat de data in het systeem zo min mogelijk fouten (datavervuiling) bevat. Een bedrijf moet op dit vlak de zaken goed op orde hebben waarbij gelet moet worden op onder andere data kwaliteit, beveiliging en structuur.

Naar aanleiding van het onderzoek is naar voren gekomen dat predictive analytics zeker in de toekomst heel belangrijk kan gaan worden voor bedrijven. Door de steeds zwaardere concurrentie zullen bedrijven het belang van predictive analytics in gaan zien waardoor de aanschaf van predictive analytics software waarschijnlijk de komende jaren toe zal gaan nemen.

De tool SAP BO Predictive Workbench zou zeer geschikt zijn om voorspellende analyses te maken in bedrijven. Zeker voor bedrijven die SAP al hebben geïmplementeerd en waarbij de data van voldoende kwaliteit is, zou dit een hele mooie toevoeging zijn. De tool biedt voor verschillende branches mogelijkheden waar veel voordeel mee kan worden gerealiseerd. Uiteindelijk kan een organisatie beslissingen gaan sturen en optimaliseren door het integreren van voorspellende technologie met als doel om de winstgevendheid te maximaliseren.

Evaluatie

Tijdens deze afstudeerstage heb ik voor het eerst in een grote organisatie gewerkt aan een project. Door vijf maanden alleen te werken aan een grotere opdracht heb ik heel veel geleerd over de werkwijze in een grote organisatie. Binnen Accenture was iedereen bereikbaar om te helpen bij eventuele problemen en daarnaast wilden zij ook graag van mijn onderzoek leren. Dit was heel erg leuk om te ervaren als stagiaire.

In het begin van mijn stage ben ik begonnen met het uitwerken van de theorie van de tool SAP BO Predictive Workbench. Hierbij duurde het even voordat ik me echt een beeld kon vormen van alle analysemodellen die ik wilde uitwerken. Het was lastig om het eenvoudig op te schrijven en niet meteen te diep in detail te treden. Bij veel van de analysemodellen waren er genoeg statistische berekeningen te vinden maar wat het model nou echt berekende bleef in het begin wat onduidelijk. Na meer bronnen te hebben gezocht en hulp van mijn begeleiders werd het steeds wat duidelijker waardoor ik er uiteindelijk toch uitkwam.

Zoals besproken is de tool na een maand geïnstalleerd en kon ik ermee gaan werken. Doordat er al demo's in de tool stonden met uitleg was het goed te doen om zelf uit te zoeken hoe alles werkte. Na een aantal demo's te hebben bestudeerd werd de werkwijze van de tool steeds duidelijker.

In november 2010 werd er een demo jam gehouden waarbij vijf medewerkers van Accenture een SAP demo konden geven over een onderwerp naar keuze. Naar aanleiding hiervan kwam het voorstel om tijdens deze bijeenkomst een presentatie te geven over mijn tool zodat iedereen kon zien waar ik mee bezig was. Aan het einde van de avond werd uit alle demo's de beste uitgekozen en dat bleek die van mij te zijn. Deze presentatie/demo heeft bijgedragen aan de kennismaking en acceptatie met SAP BO Predictive Workbench, want hierbij waren bijna alle medewerkers van de area aanwezig. De tool werd goed ontvangen en de reacties waren enthousiast, wat ook bleek uit de 1^{e} plek die ik hiermee behaald heb.

Na deze presentatie kwam het voorstel vanuit de opleiding om interviews te houden onder consultants om zo een betere onderbouwing van mijn opdracht te krijgen. Mijn begeleiders vanuit Accenture vonden dit niet meteen een goed idee omdat ze liever wilden dat ik verder ging met de tool in plaats van tijd besteden aan interviews. Hiernaast was het ook lastig om consultants te bereiken in verband met hun drukke agenda en zat waarschijnlijk niet iedereen hierop te wachten. Uiteindelijk is er besloten om een tussenweg te nemen en een mail te versturen naar een aantal consultants met vragen hierin. Hieruit bleek ook dat consultants niet echt een specifiek probleem konden aangeven waar ze bij klanten tegenaan liepen. Ze waren geïnteresseerd naar de mogelijkheden van de tool want dit was bij hun nog niet helemaal duidelijk.

Vervolgens heb ik een business case opgesteld met het onderwerp churn (weglopen van klanten) omdat dit tegenwoordig een algemeen probleem is bij bedrijven. Deze demo heb ik tevens gepresenteerd aan medewerkers die geïnteresseerd waren in de analytische tool en eventueel in de toekomst met analytische software wilde gaan werken.

Kortom, ik heb het afgelopen half jaar veel geleerd en helemaal op mijn plek gezeten bij Accenture. De opdracht was interessant en leuk om uit te voeren en ik ben uiteindelijk tot het gewenste eindresultaat gekomen.

Literatuurlijst

In deze literatuurlijst staan de bronnen weergegeven gesorteerd op nummer, waarnaar ook verwezen wordt in de scriptie (onderaan in de voetnoot).

Boeken

1. Daan van Beek, *De intelligente organisatie. Prestatieverbetering en organisatie-ontwikkeling met Business Intelligence.,* 2e druk, Tutein Nolthenius, 2008.
2. Drs. Wil Wurtz, *Retentie- en loyaliteitsprogramma's*, 1e druk, Samsom, 2000
3. Egbert Jan van Bel, *Event Driven Marketing. Op het juiste moment, met het juiste aanbod, bij de juiste klant.* 1e druk, Kluwer, 2005
4. E. Peelen, *CRM-de klant centraal.* 1e druk, Kluwer
5. Marc van de Perre & Ton Kuijlen, *Ken uw klant. Als motor van uw marketing*, 1e druk, Lannoo Camput & Scriptum, 2009.
6. Patrick de Pelsmacker & Patrick van Kenhove, *Marktonderzoek. Methoden en toepassingen.*, 2e druk, Pearson Education, 2006.
7. Pieter den Hamer, *De organisatie van business intelligence. Essentiele ICT-competenties voor managers.* 1e druk, Academic Service, 2005.
8. SPSS, *CRISP-DM 1.0.,*2000.

Websites

9. Accenture: http://www.accenture.com/
10. Computable: http://www.computable.nl
11. 12 manage – Alles over management: http://www.12manage.com/
12. SAP: http://www.sap.com/index.epx
13. SAS – Software voor business intelligence: http://www.sas.com/
14. Sentient – Krachtige datamining software: http://www.sentient.nl/?dmnl
15. SPSS - Een statistisch computerprogramma (overgenomen door IBM): http://www.spss.com/, http://www.ibm.com/
16. Wikipedia: http://www.wikipedia.nl/

Whitepapers

17. Insurance_Claims_Analytics
18. Maximizing_Revenue_and_Retention_
19. SAP_BusinessObjects_Integration_for_SPSS_Clementine_Client_Server_
20. SAP_BusinessObjects_Predictive_Workbench_
21. SAS_FraudDetection
22. SAS_CustomerAnalytics

Handleiding

23. Accenture, *SAP BI Bootcamp.,* 2006
24. M. van Kuijk, *Kennismaking SAP BW*., April 2010

Artikel

25. Bart DE MOOR, Herman VERRELST, *Fraude en fraudedetectie in de mobilofonie*., 2000

Bijlagen

Bijlage A. Project Initiatie Project (PID)

SAP BO Predictive Workbench

SAP BusinessObjects

Project Initiatie Document

Datum:	5 oktober 2010
Naam:	Sascha Maas
Studentnummer:	2098638
Afstudeerbedrijf:	Accenture
Bedrijfsbegeleiders:	Irina Mizhiritskaya Alexander Radin
Afstudeerdocent:	Ko Vleugel
Versie:	1.0
Status:	Final

Documentenhistorie

Revisies

Versie	Status	Datum	Wijzigingen
0.1	Concept	08-09-2010	Initiële uitgave
0.2	Concept	10-09-2010	Feedback bedrijfsbegeleider verwerkt
0.3	Concept	21-09-2010	Feedback afstudeerdocent verwerkt
1.0	Final	05-10-2010	Feedback afstudeerdocent verwerkt

Goedkeuring

Dit document bevat de volgende goedkeuringen:

Versie	Datum	Naam	Functie	Paraaf
0.1	08-09-2010	Irina Mizhiritskaya	Bedrijfsbegeleider	v
		Alexander Radin	Bedrijfsbegeleider	
1.0	05-10-2010	Ko Vleugel	Afstudeerdocent	v

Distributie

Dit document is verstuurd aan:

Versie	Datum	Naam	Functie
0.1	08-09-2010	Irina Mizhiritskaya	Bedrijfsbegeleider
		Alexander Radin	Bedrijfsbegeleider
0.2	10-09-2010	Ko Vleugel	Afstudeerdocent
		Irina Mizhiritskaya	Bedrijfsbegeleider
		Alexander Radin	Bedrijfsbegeleider
0.3	21-09-2010	Ko Vleugel	Afstudeerdocent
		Irina Mizhiritskaya	Bedrijfsbegeleider
		Alexander Radin	Bedrijfsbegeleider
1.0	05-10-2010	Ko Vleugel	Afstudeerdocent
		Stage- en afstudeerbureau	Secretariaat

Inhoudsopgave

1 INLEIDING 37

1.1 DOEL VAN DIT DOCUMENT 37
1.2 OPBOUW VAN DIT DOCUMENT 37

2 ACHTERGROND 38

2.1 GESCHIEDENIS 38
2.2 VESTIGINGEN 40
2.3 ORGANISATIESTRUCTUUR 41
2.4 AFSTUDEEROPDRACHT 42
2.5 AANLEIDING OPDRACHT 43

3 PROJECTDEFINITIE 44

3.1 PROJECTDOELSTELLINGEN 44
3.2 SCOPE 44
3.3 PROJECTRESULTATEN 45
3.4 GEKOZEN OPLOSSING OF AANPAK 45
3.5 UITSLUITINGEN 46
3.6 BEPERKINGEN 46
3.7 AFHANKELIJKHEDEN 46
3.8 RANDVOORWAARDEN 47
3.9 AANNAMES 47

4 PROJECTORGANISATIESTRUCTUUR 48

4.1 OPDRACHTGEVER 48
4.2 BEGELEIDERS VANUIT ORGANISATIE 48
4.3 BEGELEIDER VANUIT OPLEIDING 48
4.4 AFSTUDEERDER 48

5 PROJECTBEHEERSING 49

5.1 RAPPORTAGE 49
5.2 VOORTGANGSBEWAKING 49
5.3 RISICO- EN ISSUEMANAGEMENT 50
5.4 AFWIJKING- EN ESCALATIEPROCEDURE 50

6 BIJLAGEN 51

BIJLAGE A COMMUNICATIEPLAN 51
BIJLAGE B PROJECT KWALITEITSPLAN 52
BIJLAGE C INITIËLE BUSINESS CASE 52
BIJLAGE D INITIEEL PROJECTPLAN 53
BIJLAGE E INITIEEL RISICOLOGBOEK 55
BIJLAGE F ORGANOGRAM BEDRIJFSMANAGEMENT 56

1 Inleiding

Mijn afstudeerstage bestaat uit een opdracht bij Accenture in Den Bosch.

Voor deze opdracht ga ik onderzoek doen naar een nieuwe tool: 'Predictive Workbench'. Predictive Workbench is krachtige software voor geavanceerde voorspellende analyses. Het resultaat van de combinatie van kennis en geavanceerde analyses is een helder toekomstbeeld. Zo wordt er bijvoorbeeld inzicht verkregen op het te verwachten gedrag van klanten.

Door middel van deze analyses kunnen de prestaties van een onderneming verbeterd worden.

Het doel van de opdracht is om Predictive Workbench te beheersen en vervolgens te kunnen toepassen op de vooraf opgestelde business case.

1.1 Doel van dit document

Het doel van dit document is om de afstudeeropdracht te definiëren, waardoor het als basis kan dienen voor het verloop van de afstudeerstage.

Het document beschrijft wat de opdracht inhoudt en welke afspraken er zijn gemaakt. Zo wordt precies beschreven wat wel en niet binnen de scope valt. Hiernaast biedt het document duidelijkheid over de doelstellingen en de producten van de opdracht.

Ook wordt in dit document weergegeven wie bij de opdracht zijn betrokken en wat de taken van deze betrokkenen zijn.

1.2 Opbouw van dit document

In hoofdstuk 2 wordt de achtergrond van Accenture beschreven. Hierin wordt ingegaan op de geschiedenis, de vestigingen, organisatiestructuur en de opdracht.

Hoofdstuk 3 bevat een uitleg over afstudeeropdracht in de projectdefinitie. Hierbij wordt verder ingegaan op doelstellingen, scope, projectresultaten, gekozen oplossing of aanpak, uitsluitingen, beperkingen, afhankelijkheden, randvoorwaarden en aannames.

In hoofdstuk 4 wordt ingegaan op de projectorganisatiestructuur. Hierin worden alle betrokkenen van de opdracht beschreven.

In hoofdstuk 5 bevat een uitleg op de invoering en handhaving van een stelsel van maatregelen en procedures. Hierbij wordt de beheersing beschreven van de rapportage, voortgangsbewaking, risico- en issuemanagement en op de afwijking- en escalatieprocedure.

En als laatste is in de bijlagen het communicatieplan, kwaliteitsplan, de initiële business case en het organogram van het bedrijfsmanagement toegevoegd.

2 Achtergrond

Accenture is een internationaal consultancy bedrijf dat diensten aanbiedt voor consulting, technologie en outsourcing.

Consulting heeft als doel de bedrijfsresultaten te optimaliseren. Accenture heeft een diepgaande betrokkenheid bij een organisatie met strategische doelen en zorgt voor effectieve oplossingen. Hierbij kiest Accenture niet alleen voor een traditionele benadering, maar juist dankzij de breedte van de dienstverlening ook voor het ontwikkelen, implementeren en beheren van innovatieve IT-gerelateerde oplossingen.

Bij technologie worden business oplossingen van onder meer SAP, Siebel, People Soft en Oracle ontwikkeld, getest en geïmplementeerd. Accenture vertaalt innovatieve technologie in innovatieve oplossingen. De wisselwerking tussen Technologie en Consulting is sterk en biedt zowel voor Accenture als voor opdrachtgevers interessante kansen en mogelijkheden.

Outsourcing diensten gaan van applicatiemanagement tot optimalisatie van bedrijfsprocessen en infrastructuur outsourcing. Het wordt ingezet als een onmisbaar instrument om de prestaties van klanten te verbeteren.

2.1 Geschiedenis

Accenture heeft een bewogen historie, die meer dan een halve eeuw teruggaat. Van de eerste dagen als pionier in een nieuwe wereld van informatietechnologie tot de positie vandaag de dag als marktleider op het gebied van Consulting, Technologie en Outsourcing.

2.1.1 De start

Het begon in 1913, toen de Amerikaan Arthur Edward Andersen samen met een partner besloot Arthur Andersen & Co op te richten. Tot aan zijn dood in 1947 was Andersen als senior partner nauw verbonden met de ontwikkeling en groei van zijn bedrijf.

In deze jaren 40 boden de recent uitgevonden computers de mogelijkheid voor bedrijven om hun productiviteit aanzienlijk te verhogen. Er waren echter maar weinig captains of industry die het volledig potentieel hiervan erkenden. Visionaire leiders van Arthur Andersen als Leonard Spacek, John Higgings en Joe Glickauf waren de uitzonderingen. Zij gaven de computers een plaats in de dagelijkse bedrijfsvoering en ontwikkelden zo een belangrijke rol voor consultants in de introductie van computertechnologie bij andere bedrijven.

Afbeelding 14 – Joe Glickauf

Technologische innovatie ligt diep geworteld in de cultuur van Accenture.

2.1.2 Consulting & outsourcing

Klanten begonnen zich steeds meer te richten op het inzetten van computertechnologie om effectiever en efficiënter te gaan werken. Zodoende begon ook consulting binnen Arthur Andersen een steeds belangrijkere rol in te nemen en groeide het bedrijf uit tot een leidend consultancy bedrijf op wereldniveau.

In 1989 vormde een aantal partners van deze divisie onder de naam Andersen Consulting een nieuwe organisatie, waarbij de focus lag op Consulting en Technology Services. In 1991 kwam hier ook Outsourcing bij, toen Andersen Consulting haar eerste grote outsourcing project startte door accounting, finance en andere functies uit te gaan voeren. De wereldwijde verspreiding van technologie zorgde voor een competitieve druk op bedrijven in verschillende industrieën. Andersen Consulting begon met het aanbieden van geïntegreerde bedrijfsoplossingen op gebied van consulting, technology en outsourcing in verschillende industrieën.

2.1.3 Accent op de toekomst

In de jaren daarna ontstond een steeds grotere afstand tussen beide organisaties. Het mondde uiteindelijk uit in een definitieve scheiding in 2001. Andersen Consulting ging verder onder de naam Accenture, een samentrekking van 'Accent' on the 'Future'. Deze vondst van een Noorse medewerker symboliseerde op compacte wijze de drang om een internationale voorhoedespeler te zijn.

Afbeelding 15 – Think global, act local

Accenture heeft zich vanaf 2001 explosief ontwikkeld. Het is in Nederland tegen de economische ontwikkelingen in met gemiddeld meer dan 20 procent per jaar gegroeid. Accenture heeft een internationale geschiedenis en op dit moment een netwerk van vestigingen in 49 landen. Deze wereldwijde slagkracht combineert het met een solide plaatselijke verankering. 'Think global, act local' is een motto dat ook Accenture Nederland ver heeft gebracht.

2.2 Vestigingen

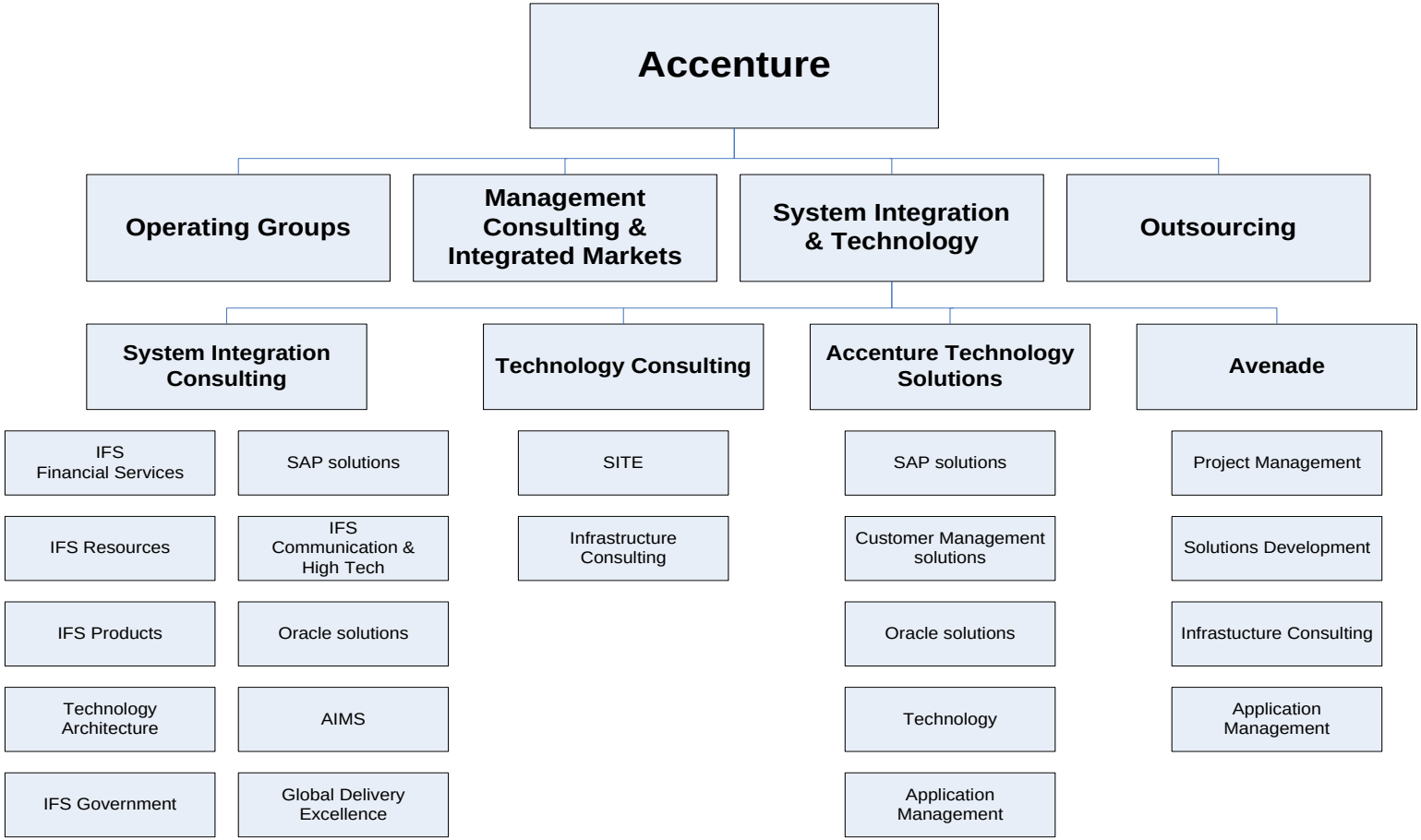
Accenture is een internationale organisatie met 110 kantoren in 49 landen. In de afbeelding hieronder staan een aantal landen weergegeven die hiertoe behoren. De hoofdvestiging in Nederland is in Amsterdam. Hiernaast zijn er ook nog vestigingen in Den Bosch en Almere.

Afbeelding 16 - Vestigingen Accenture wereldwijd

2.3 Organisatiestructuur

In afbeelding 17 is het organisatieschema van Accenture weergegeven. In dit schema staat een gedetailleerdere versie van 'System Integration & Technology'. Voor het algemene organisatieschema zie bijlage F.

De opdracht zal worden uitgevoerd bij de afdeling: 'Accenture Technology Solutions (ATS)', onder het cluster: 'SAP solutions'.



Afbeelding 17 - Organisatiestructuur Accenture

2.4 Afstudeeropdracht

Voor de afstudeeropdracht wordt gebruik gemaakt van de Business Intelligence (BI) oplossingen van SAP. SAP is wereldwijd marktleider in softwareoplossingen voor zowel kleine en middelgrote bedrijven als internationale ondernemingen. De doelstelling van Business Intelligence is de juiste mensen, op het juiste moment en de juiste plaats van de juiste informatie te voorzien. Aan de hand van deze informatie kunnen tijdige beslissingen worden gemaakt.

2.4.1 SAP

De oplossingen van SAP helpen organisaties hun bedrijfsvoering te stroomlijnen en aanzienlijke efficiencyverbeteringen te bereiken binnen hun toeleveringsketens en bedrijfsprocessen. SAP heeft meer dan 25 branchespecifieke oplossingen voor markten variërend van high tech industrie en retail tot overheids- en financiële instellingen.

Het belangrijkste product van SAP is SAP Enterprise Resource Planning (ERP). Met een ERP-pakket kunnen in principe alle bedrijfsfuncties worden ondersteund.
SAP is een informatie- en besturingssysteem waarin bedrijfsmatige processen kunnen worden vastgelegd en beheerd. Deze processen zijn ondergebracht in modules. De gegevens in de modules worden onderling direct uitgewisseld, waardoor een volledig geïntegreerd systeem ontstaat.

Binnen SAP wordt alle informatie in een relationele structuur vastgelegd, dat gebaseerd is op een business model. Dit kan voor elke klant apart worden geconfigureerd. Deze structuur vindt zijn weerslag in de opbouw van de databases. In de SAP-structuur kunnen zeer grote organisaties worden vastgelegd met onderverdeling in divisies, verkooporganisaties, fabrieken, vestigingen, magazijnen, enzovoorts.

2.4.2 Business Intelligence

De opdracht is gericht op de Business Intelligence oplossingen van SAP. Business Intelligence staat voor het verzamelen van informatie binnen de eigen handelsactiviteit. Het kan omschreven worden als het proces van gegevens omzetten in informatie, die vervolgens leidt tot kennis.

Betrouwbare, tijdige informatie is essentieel voor de bedrijfsvoering. Of het nu gaat om de concurrentiepositie, meer grip op bedrijfsprocessen of het reduceren van kosten. Het snel kunnen beschikken over informatie is de sleutel tot het bereiken van strategische doelen.

2.5 Aanleiding Opdracht

Accenture heeft bij een groot aantal klanten Business Intelligence oplossingen geïmplementeerd met behulp van SAP BI.
SAP heeft recentelijk een aantal overnames gedaan. Een van deze overnames is die van BusinessObjects (BO).

2.5.1 BusinessObjects

SAP BusinessObjects omvat een groot aantal tools en toepassingen waarmee de prestaties van een onderneming kunnen worden verbeterd door mensen, gegevens en bedrijven via bedrijfsnetwerken te koppelen.

De Business Intelligence oplossingen van SAP BusinessObjects voorzien de gebruiker van tools die ze nodig hebben om zinvolle en doordachte beslissingen te baseren op solide gegevens en analyses. Hierdoor wordt iedereen, van de topanalist tot de incidentele zakelijke gebruiker, van de benodigde informatie voorzien, zonder al te veel afhankelijkheid van IT-afdelingen en ontwikkelaars.

Dankzij deze krachtige oplossingen kunnen eindgebruikers overal in de onderneming data bekijken, indelen, analyseren, verkennen en delen met collega's afgestemd op functie.

2.5.2 Predictive Workbench

Een onderdeel van SAP BO is Predictive Workbench, een krachtige software voor geavanceerde voorspellende analyses. In SAP BO is Predictive Workbench gebaseerd op PASW Modeler van SPSS, een van de meest krachtige en bekende tools waarmee bedrijfsanalisten data analyseren. De software kan voor deze analyses bovendien actuele gegevens uit talloze externe bronnen betrekken.

Predictive Workbench is een nieuwe geïntegreerde tool binnen SAP die tijdens de afstudeeropdracht voor het eerst bij Accenture wordt toegepast.
In bedrijven worden er wel veel analyses gemaakt maar voorspellingen worden hier nog niet vaak mee gedaan. Door middel van deze tool zouden bedrijfsanalisten (in plaats van wiskundige) voorspellende analyses kunnen maken.

Types analyse die hiermee kunnen worden uitgevoerd, omvatten onder andere het ontdekken van trends en het voorspellen van risico’s. Integratie met SAP BO staat toe om deze analyseresultaten snel en gemakkelijk te visualiseren, bijvoorbeeld in dashboards. Hierin kunnen dynamische tabellen en grafieken worden weergegeven.

Tijdens de opdracht zal onderzoek gedaan worden naar de Predictive Workbench in samenwerking met Business Intelligence. Hierbij zal worden gekeken welke inspanningen nodig zijn voor het implementeren van de tool binnen organisaties en hoe deze het beste ingezet kan worden bij klanten van Accenture.

3 Projectdefinitie

Dit hoofdstuk geeft de essenties en definitie van de afstudeeropdracht weer.

3.1 Projectdoelstellingen

Gedurende de afstudeerperiode moet er worden gewerkt met de tool SAP BO Predictive Workbench. De doelstellingen die hiervan zijn:

- Theorie Predictive Analytics beheersen
- De tool Predictive Workbench dusdanig inhoudelijk beheersen dat op basis van de opgedane kennis een inschatting gemaakt kan worden van hoe het pakket bij klanten geïmplementeerd kan worden.
    - Daarbij moet een oordeel gegeven kunnen worden over de uitvoering van verschillende typen analyses;
- Bruikbare business case opstellen
    - Predictive Workbench hierop toepassen

3.2 Scope

Onder de scope van de opdracht vallen de onderwerpen:

- Predictive Analytics
- SAP BO Predictive Workbench

Bij Predictive Analytics wordt alleen de theorie uitgewerkt maar bij SAP BO Predictive Workbench valt ook het beheersen van de tool onder de scope. Voor deze tool wordt ook een koppeling gemaakt met de brongegevens van BusinessObjects. Aan de hand van deze brongegevens worden de voorspellende analyses uitgevoerd in Predictive Workbench. In Predictive Workbench worden ook de resultaten weergeven in tabellen/grafieken. Aan de hand van deze analyses kunnen conclusies getrokken worden.
Naast deze onderwerpen wordt een business case uitgewerkt waarop vervolgens SAP BO Predictive Workbench wordt toegepast.

3.3 Projectresultaten

Het eindresultaat van de opdracht bestaat uit de volgende documenten:

- Project Initiatie Document
- Tussenrapport Predictive Analytics
    - Beschrijving modellen
        - Presentatie
    - Beschrijving verschillende soorten toepassende analyses
        - Fraude
        - Retentieanalyse
        - Klantanalyse
        - Risicoanalyse/management
- Tussenrapport Predictive Workbench
    - Beschrijving mogelijkheden tool
        - Beschikbare bronsystemen
- Business Case
    - Uitwerken business case
        - Omschrijving bedrijf
        - Business vraag opstellen
        - Methode en onderwerp bepalen
        - Model bepalen
        - Testgegevens opstellen
        - Demo geven
        - Conclusie en advies opstellen
- Scriptie

3.4 Gekozen oplossing of aanpak

Als eerste zal onderzoek worden gedaan naar het onderwerp Predictive Analytics. Hiervoor zullen verschillende analyse methodes worden uitgewerkt.
Vervolgens zal informatie worden gezocht naar Predictive Workbench en aan de hand hiervan zal de tool worden uitgeprobeerd.
Wanneer de basis van de tool wordt beheerst kan een Business case worden opgesteld. Deze Business Case wordt opgesteld aan de hand van voorbeelden van eerder opgestelde Business Cases binnen Accenture. Aan de hand van deze Business Case wordt één analyse methode uitgewerkt en toegepast op de bedrijfsgegevens van een fictief bedrijf. De gekozen analyse methode in de Business Case wordt toegepast in Predictive Workbench.

3.5 Uitsluitingen

Voor de afstudeeropdracht gelden een aantal uitsluitingen die niet tot de opdracht gerekend worden:

- Pakketselectie voor methoden van data analyse;
- Diepgaande wiskundige analyses;
- Integratie met SAP BO om de analyseresultaten snel en gemakkelijk te kunnen visualiseren;

Mocht er eventueel tijd over zijn in deze afstudeerperiode zou dit laatste puntje wel als extra kunnen worden uitgevoerd.

3.6 Beperkingen

Voor deze opdracht gelden een aantal voorwaarden waarmee rekening gehouden moet worden gedurende de afstudeerperioden:

- Het PID moet goedgekeurd zijn voor het eerste bedrijfsbezoek;
- In week 10 is er een terugkomstdag gepland;
    - Hiervoor moet een tussentijdsrapportage worden ingeleverd;
- Eén week voor de definitieve inleverdatum van de scriptie moet een concept versie worden ingeleverd bij de afstudeerbegeleider vanuit Fontys;
- De Scriptie moet 13 januari 2011 ingeleverd zijn;

3.7 Afhankelijkheden

Het onderzoek is afhankelijk van:

- Externe partijen wat betreft kennis over de werkwijze met SAP BO Predictive Workbench, mochten er eventuele moeilijkheden ontstaan;
- Toegang tot benodigde software;

3.8 Randvoorwaarden

Om tot het juiste resultaat binnen de gestelde tijd te komen, zijn hier randvoorwaarden aan gekoppeld. De voorwaarden betreffen acties die moeten worden genomen of situaties die aan bepaalde eisen moeten voldoen.
De huidige voorwaarden betreffen:

- Goede werkplek;
    - Voorzien van Laptop;
        - Met toegang tot de gebruikelijke zaken als Internet, E-mail, Office;
        - Toegang tot SAP BO Predictive Workbench;
- Goede begeleiding
    - Vanuit de opleiding en het bedrijf;
    - Mogelijkheid tot overlegsituaties om de voortgang van de afstudeeropdracht te bespreken;
    - Feedback krijgen op ingeleverde verslagen;

3.9 Aannames

De aannames die worden gedaan aan de hand van de afstudeeropdracht:

- SAP BO Predictive Workbench is een pakket dat zonder uitgebreide cursus/training te beheersen is;
- Dat medewerking wordt verleend wanneer dit nodig blijkt te zijn door:
    - Fontys met betrekking tot de scriptie en het afstudeertraject;
    - Accenture voor het verloop van de opdracht;

4 Projectorganisatiestructuur

Dit hoofdstuk beschrijft de opzet en invulling van de projectorganisatie.

4.1 Opdrachtgever

Bedrijfsnaam:	Accenture
Bezoekadres:	Hambakenwetering 5K, 2e verdieping
Postcode en plaats:	5231 DD Den Bosch
Telefoon:	+31 20 593 9900
Internet:	www.accenture.com
Afdeling:	ATS (BI)

4.2 Begeleiders vanuit organisatie

Naam:	Irina Mizhiritskaya
Afdeling:	Solutions/BI
Telefoonnummer:	+31 20 5729297
E-mail adres:	irina.mizhiritskaya@accenture.com

Naam:	Alexander Radin
Afdeling:	Solutions/BI
Telefoonnummer:	+31 20 5729353
E-mail adres:	alexander.radin@accenture.com

4.3 Begeleider vanuit opleiding

Naam:	Ko Vleugel
Organisatie onderdeel:	Fontys Hogescholen ICT
Gebouw:	Eindhoven, Rachelsmolen 1 (R1)
Kamer:	4.07
Telefoonnummer:	0877 870988
E-mail adres stage:	k.vleugel@fontys.nl

4.4 Afstudeerder

Naam:	Sascha Maas
Adres:	Weisshorn 175
Postcode:	5624 NW Eindhoven
Telefoonnummer:	06-42454421
E-mail adres stage:	sascha.maas@accenture.com
E-mail adres privé:	maas.sascha@gmail.com

5 Projectbeheersing

5.1 Rapportage

In onderstaande matrix staat aangegeven welke rapportages naar welke betrokkenen gestuurd moeten worden. Daarnaast wordt aangegeven door wie, welke rapporten goedgekeurd moeten worden of wanneer deze alleen ter informatie ontvangen worden.

Rapport	Stage coördinator	Begeleider Fontys	Begeleiders Accenture	Afstudeerder
Projectvoorstel	A + G + D	I	O	O + D
PID		T + A + G	A + G	O + D
Logboek		T + A + G		O + D
Scriptie		A + T + G + D	A + G	O + D

Legenda:

O	Opstellen	**A**	Adviseren	**I**	Ontvangen ter informatie
T	Toetsen	**D**	Distribueren/archiveren	**G**	Goedkeuren

Tabel 8 - Rapportage matrix

5.2 Voortgangsbewaking

In onderstaande matrix staat weergeven hoe en met welke frequentie de voortgang wordt bijgehouden de komende 5 maanden.

Overleg	Aanwezig	Frequentie	Tijdstip	Doel
Voortgang opdracht	Alexander Radin Irina Mizhiritskaya Sascha Maas	1 x per week	Vrijdag 10 uur	Voortgang bepalen
Voortgang afstuderen	Charlotte Hoffman Sascha Maas	1 x per maand	Nader te bepalen	Voortgang bepalen
Bedrijfsbezoek	Ko Vleugel Alexander Radin Irina Mizhiritskaya Sascha Maas	2 x per stageperiode	Nader te bepalen	Voortgang bespreken

Tabel 9 - Voortgangsbewaking matrix

5.3 Risico- en issuemanagement

Door de wekelijkse besprekingen met de begeleiders van Accenture (zie paragraaf 5.2) kunnen mogelijke issues en risico's tijdig worden gesignaleerd en opgepakt.
Doordat de begeleiders al veel ervaring en affiniteit hebben met het onderwerp van de opdracht, kunnen eventuele issues worden besproken en mogelijk worden opgelost.

Wanneer dit toch niet mocht lukken, dan is binnen de afdeling van het bedrijf veel kennis aanwezig over het onderwerp van de opdracht. Er zijn voldoende mogelijkheden binnen de afdeling om maatregelen tegen mogelijke risico's en issues te nemen.

In bijlage E is een initieel risicologboek opgesteld waarin mogelijke risico's vermeld worden met bijbehorende maatregelen om de gevolgen zo veel mogelijk te beperken.

5.4 Afwijking- en escalatieprocedure

Als blijkt dat afwijkingen ontstaan op de huidige planning, wordt dit overlegd in een bijeenkomst met de begeleiders. Hierin wordt besproken hoe deze afwijking tot stand heeft kunnen komen en wat het gevolg hiervan is. Verder wordt gekeken hoe er met de afwijking om moet worden gegaan. Wanneer dit bekend is kan de planning worden aangepast waardoor weer verder kan worden gegaan met de opdracht aan de hand van een corresponderende planning.

6 Bijlagen

Bijlage A Communicatieplan

Communicatiekanalen

Gedurende de afstudeerperiode zullen onderstaande communicatiekanalen voorkomen.

Van	Naar	Informatie	Medium	Frequentie van data
Afstudeerder	Opleidingsbegeleider	Logboek	E-mail	Wekelijks
Opleidingsbegeleider	Afstudeerder	Terugkoppeling	E-mail	Wekelijks
Afstudeerder	Opleidingsbegeleider	PID	E-mail	Tot goedkeuring
Opleidingsbegeleider	Afstudeerder	Terugkoppeling PID	E-mail	Tot voldoende resultaat
Afstudeerder	Opleidingsbegeleider	Tussentijdse rapportage	E-mail	Eenmalig
Opleidingsbegeleider	Afstudeerder	Terugkoppeling tussentijdse rapportage	Mondeling	Eenmalig
Afstudeerder	Opleidingsbegeleider	Concept scriptie	E-mail	Eenmalig (indien nodig meer)
Opleidingsbegeleider	Afstudeerder	Terugkoppeling concept scriptie	E-mail	Eenmalig (indien nodig meer)
Afstudeerder	Bedrijfsbegeleider	Voortgang	Mondeling / E-mail	Wekelijks
Bedrijfsbegeleider	Afstudeerder	Terugkoppeling	Mondeling / E-mail	Wekelijks
Afstudeerder	Opleidingsbegeleider	Bedrijfsbezoek	Mondeling	2 x per stageperiode
Afstudeerder	Opleidingsbegeleider	Afstudeerzitting	Mondeling	Eenmalig

Tabel 10 - Communicatiekanalen matrix

Incidenteel

In de onderstaande tabel staan de communicatiekanalen die niet stelselmatig zullen voorkomen.

Van	Naar	Informatie	Medium	Situatie
Afstudeerder	Bedrijfsbegeleider	Afwezigheid	E-mail / telefoon	Absentie
Afstudeerder	Bedrijfsbegeleider	Aanvraag overleg	E-mail / telefoon	Onvoorziene omstandigheden
Afstudeerder	Opleidingsbegeleider	Aanvraag overleg	E-mail / telefoon	Onvoorziene omstandigheden
Bedrijfsbegeleider	Afstudeerder	Aanvraag overleg	E-mail / telefoon	Onvoorziene omstandigheden
Opleidingsbegeleider	Afstudeerder	Aanvraag overleg	E-mail / telefoon	Onvoorziene omstandigheden

Tabel 11 - Incidentele communicatie matrix

Bijlage B Project kwaliteitsplan

Gedurende de opdracht wordt de kwaliteit van de werkzaamheden bewaakt om tot een kwalitatief goed eindresultaat te komen. Door middel van frequente terugkoppelingsmomenten, kan de kwaliteit van de opdracht goed worden bijgehouden.

Het resultaat van de opdracht moet aan de volgende eisen voldoen:

- De planning is gevolgd;
- Onderzoek heeft alle doelstellingen (geformuleerd in paragraaf 3.1) behaald;
- Business Case is opgesteld aan de hand van voorgaande voorbeelden binnen Accenture;
- Inhoud is voor de lezer goed te begrijpen;
- De conclusie geeft antwoord op de vraag hoe Predictive Workbench het beste kan worden ingezet bij klanten;

Bijlage C Initiële Business Case

Kosten

De kosten die op dit moment bekend zijn bestaan uit de stagevergoeding per maand voor de afstudeerder. Deze kosten zullen uitkomen op:

- 5 Maanden à € 425,- = € 2.125,-

Baten

De baten van deze opdracht zullen bestaan uit onderstaande onderwerpen:

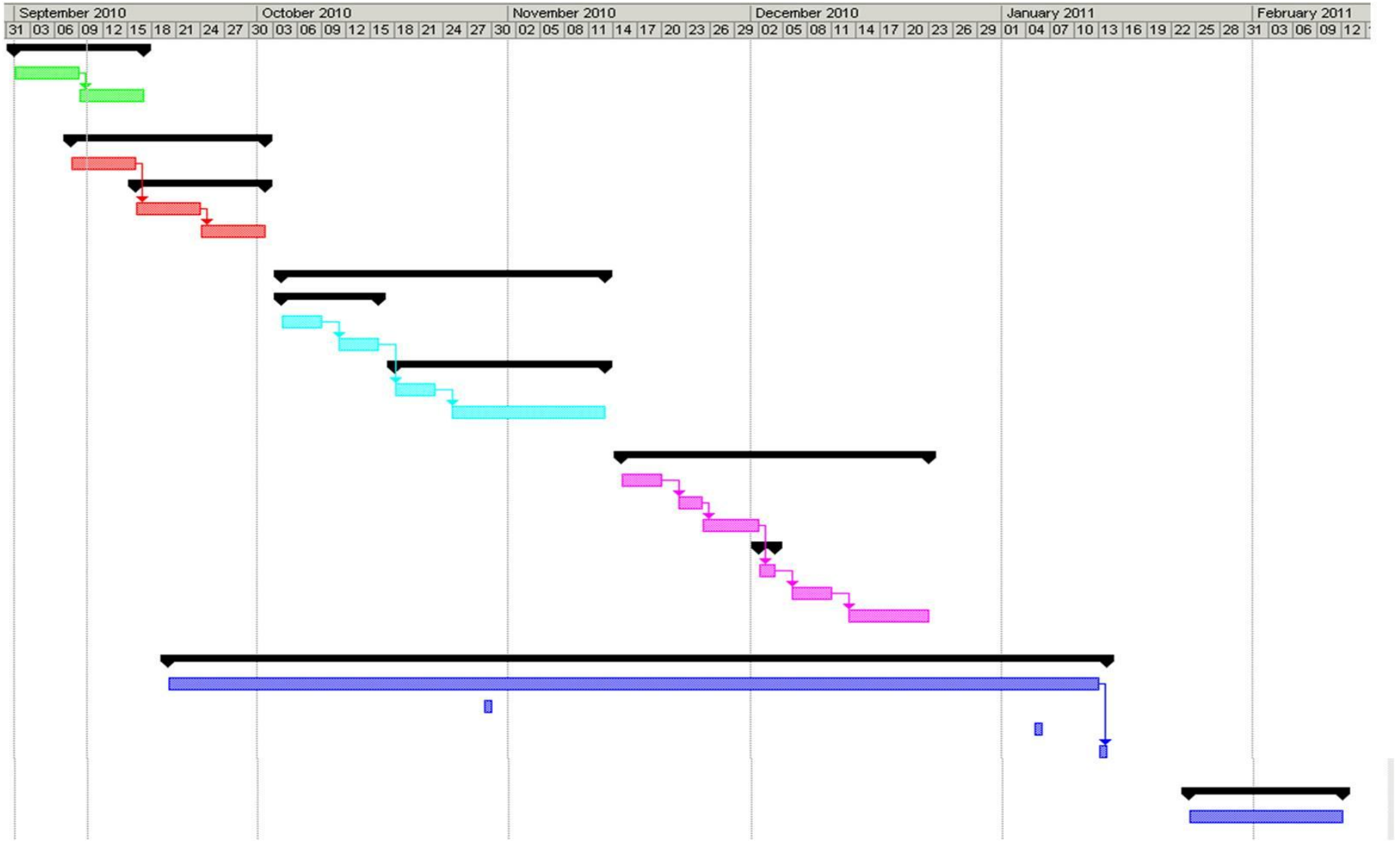
- Inzicht in de werking van Predictive Workbench;
- Verschillende mogelijkheden voor analyses die men door middel van Predictive Workbench kan uitvoeren;
- Inzicht in de inspanning die nodig is om Predictive Workbench te implementeren;
- Inzicht in hoe Predictive Workbench het beste ingezet kan worden bij klanten;

Bijlage D Initieel Projectplan

In de onderstaande afbeeldingen is de tijdsplanning voor deze afstudeeropdracht ingevoerd in Microsoft Project. Eerst is de activiteiten planning weergeven en vervolgens de tijdlijn planning.

	Task Name	Duration	Start	Finish
	Project Initiatie Document	**12 days**	**Wed 01-09-10**	**Thu 16-09-10**
	Concept PID schrijven	6 days	Wed 01-09-10	Wed 08-09-10
	Definitief PID uitwerken	6 days	Thu 09-09-10	Thu 16-09-10
	Predictive analystics	**18 days**	**Wed 08-09-10**	**Fri 01-10-10**
	Theorie bestuderen	6 days	Wed 08-09-10	Wed 15-09-10
	Theorie uitwerken	**12 days**	**Thu 16-09-10**	**Fri 01-10-10**
	Mogelijke methodes uitwerken	6 days	Thu 16-09-10	Thu 23-09-10
	Methodes koppelen aan verschillende branches	6 days	Fri 24-09-10	Fri 01-10-10
	Predictive Workbench	**30 days**	**Mon 04-10-10**	**Fri 12-11-10**
	Beschrijving mogelijkheden	**10 days**	**Mon 04-10-10**	**Fri 15-10-10**
	Beschrijving Predictive Workbench	5 days	Mon 04-10-10	Fri 08-10-10
	Beschrijving beschikbare bronsystemen	5 days	Mon 11-10-10	Fri 15-10-10
	Kennismaking met Predictive Workbench	**20 days**	**Mon 18-10-10**	**Fri 12-11-10**
	Mogelijkheden Predictive Workbench ontdekken	5 days	Mon 18-10-10	Fri 22-10-10
	Oefenen met bestaande trainingen	15 days	Mon 25-10-10	Fri 12-11-10
	Business case	**28 days**	**Mon 15-11-10**	**Wed 22-12-10**
	Opstellen business case	5 days	Mon 15-11-10	Fri 19-11-10
	Methode en onderwerp bepalen	3 days	Mon 22-11-10	Wed 24-11-10
	Business vraag bepalen	5 days	Thu 25-11-10	Wed 01-12-10
	Model bepalen	**2 days**	**Thu 02-12-10**	**Fri 03-12-10**
	Onderzoeken welke systemen je moet gebruiken	2 days	Thu 02-12-10	Fri 03-12-10
	Inrichten testomgeving	5 days	Mon 06-12-10	Fri 10-12-10
	Analyses uitvoeren	8 days	Mon 13-12-10	Wed 22-12-10
	Rapportage	**84 days**	**Mon 20-09-10**	**Thu 13-01-11**
	Scriptie schrijven	83 days	Mon 20-09-10	Wed 12-01-11
	Tussentijdse scriptie inleveren	1 day	Fri 29-10-10	Fri 29-10-10
	Concept scriptie inleveren	1 day	Wed 05-01-11	Wed 05-01-11
	Definitieve scriptie inleveren	1 day	Thu 13-01-11	Thu 13-01-11
	Afstudeerzittingen	**15 days**	**Mon 24-01-11**	**Fri 11-02-11**
	Afstudeerzitting	15 days	Mon 24-01-11	Fri 11-02-11

Tabel 12 – Activiteiten planning



Afbeelding 18 - Tijdlijn planning

Bijlage E Initieel Risicologboek

In het onderstaande risicologboek zijn alle mogelijke projectrisico's geregistreerd.

Nr.	Omschrijving	Kans[13]	Effect[14]	Risico[15]	Wanneer
1	Wijziging planning	0,7	2	1,4	Gehele project
2	Doelstellingen niet halen	0,2	4	0,8	Gehele project
3	Wijziging eisen	0,3	3	0,9	Gehele project
4	Onvoldoende kennis bij de afstudeerder	0,1	3	0,3	Gehele project
5	Vastlopen bij de opdracht	0,5	3	1,5	Gehele project
6	SAP BO Predictive Workbench niet op tijd beschikbaar	0,3	4	1,2	Begin van project

Tabel 13 - Projectrisico's

De maatregelen die tegen bovenstaande risico's worden genomen, staan in onderstaande tabel weergeven.

Nr.	Maatregelen
1	Vanaf het begin de planning goed bijhouden en mocht hiervan afgeweken worden de planning meteen aanpassen. Hierdoor kan er precies worden gekeken hoeveel tijd er nog voor welk onderdeel staat. Dit zorgt voor inzicht in de werkzaamheden, waardoor de planning alsnog gehaald kan worden.
2	Vanaf het begin van de afstudeerperiode de doelstellingen goed formuleren. Wanneer deze doelstellingen duidelijk zijn kan er wekelijks worden gekeken in hoeverre alle doelstellingen zijn gerealiseerd en er rekening mee worden gehouden.
3	Gedurende het project kunnen eisen altijd wijzigen. Een flexibele houding is hierbij belangrijk maar er moet ook goed worden gekeken wat het nut en de noodzaak van die betreffende wijzigingen is.
4	Wanneer het blijkt dat de afstudeerder onvoldoende kennis heeft om deze afstudeeropdracht uit te kunnen voeren, zal dit door middel van zelfstudie moeten worden opgelost. Eventueel kan er ook worden gekeken naar een persoon die deze extra kennis bij kan brengen.
5	Er kan advies over de opdracht worden gevraagd aan de begeleiders, docenten of andere personen die kennis hebben van het onderwerp. Hiernaast kan er ook worden gekeken of deze informatie in boeken, artikelen of op internet te vinden is.
6	Vanwege licentie redenen is het de vraag of de tool op tijd beschikbaar is. Tot die tijd zal de theorie over Predictive Analytics worden uitgewerkt zodat dit deel van de opdracht klaar is en de benodigde kennis aanwezig is, wanneer er kan worden gestart met SAP BO Predictive Workbench. Hiernaast zal er regelmatig worden geïnformeerd naar de tool en de datum waarop deze beschikbaar is voor gebruik.

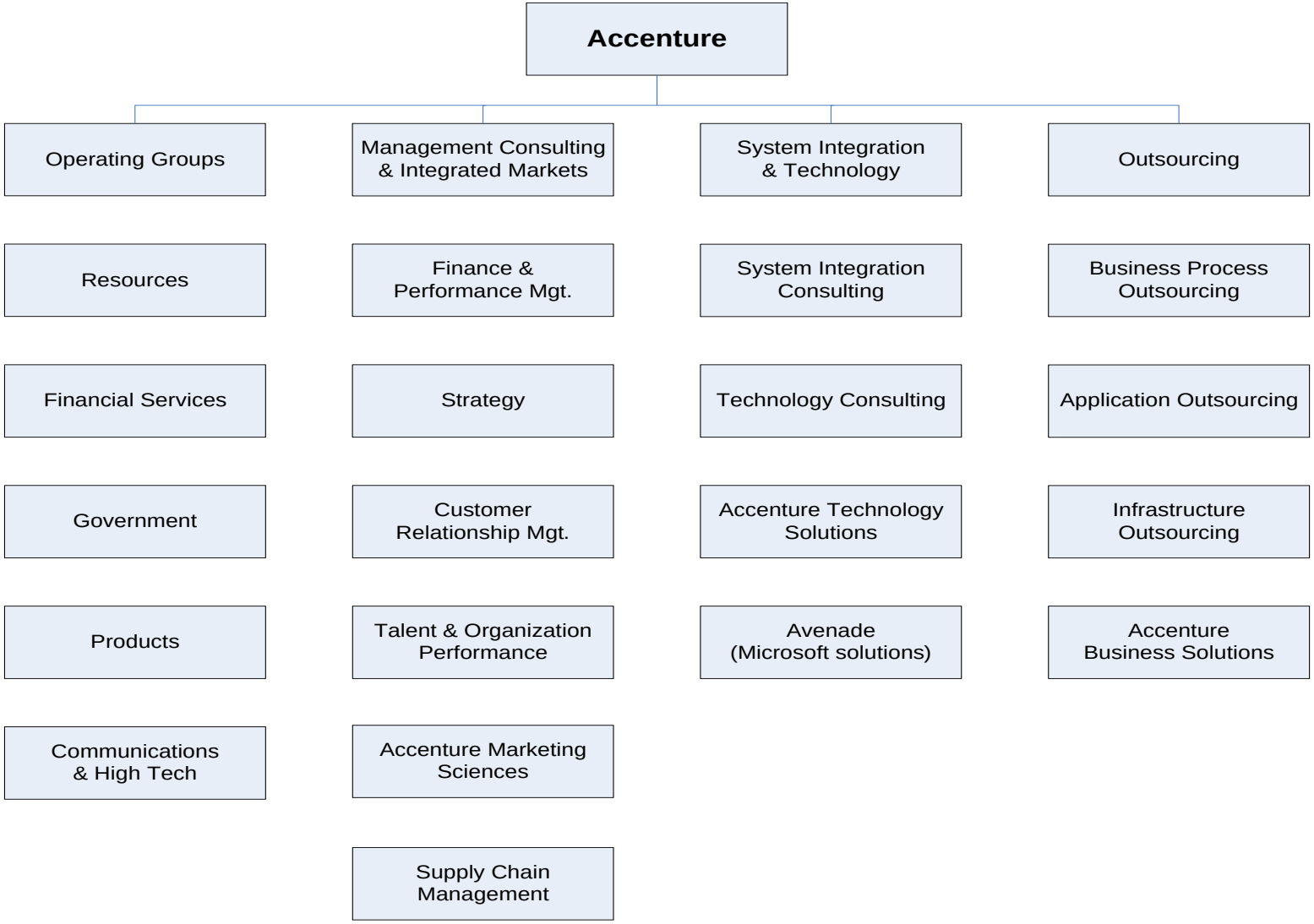
Tabel 14 - Maatregelen Projectrisico's

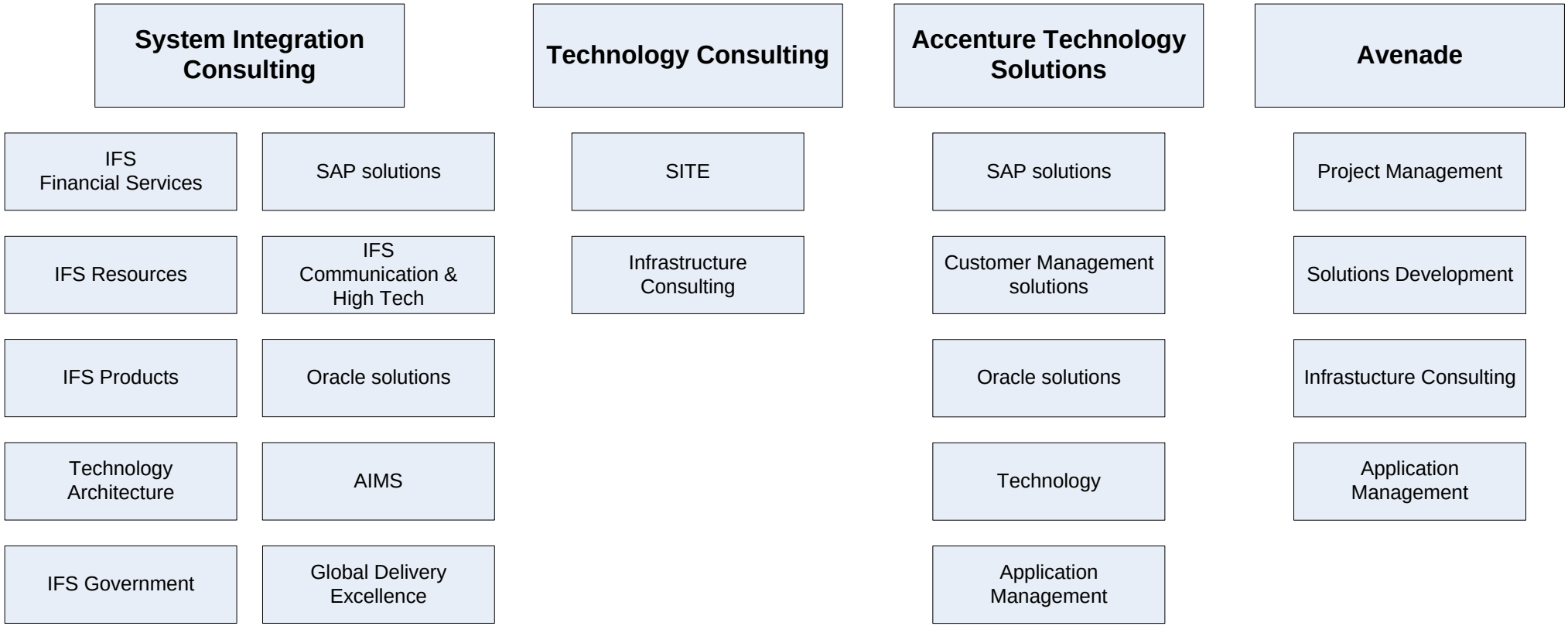
[13] *Kans - Kans van optreden wordt geschat tussen de 0 en 1.*

[14] *Effect - Het effect wordt ingeschat als 1 (gering), 2 (laag), 3 (middel) of 4 (hoog).*

[15] *Risico - Het risico wordt berekend door het product te nemen van de kans en het effect.*

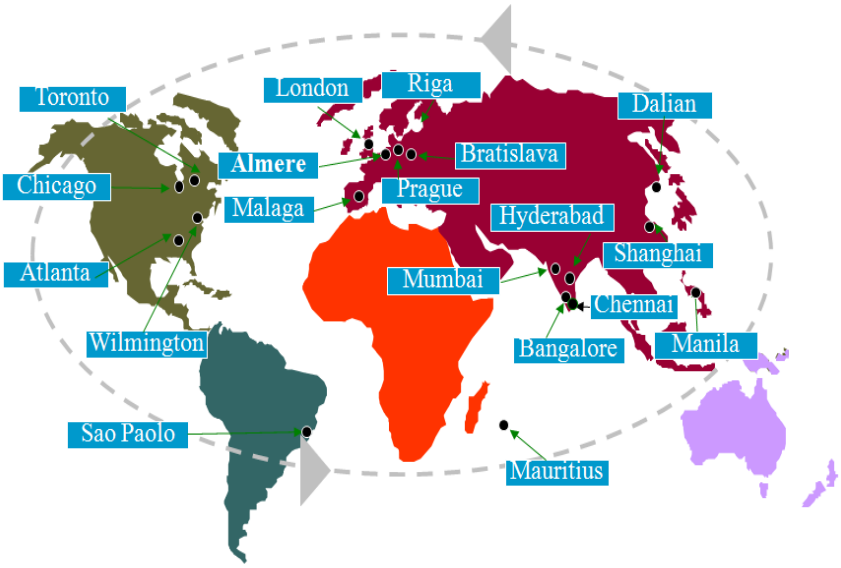
Bijlage F Organogram Bedrijfsmanagement





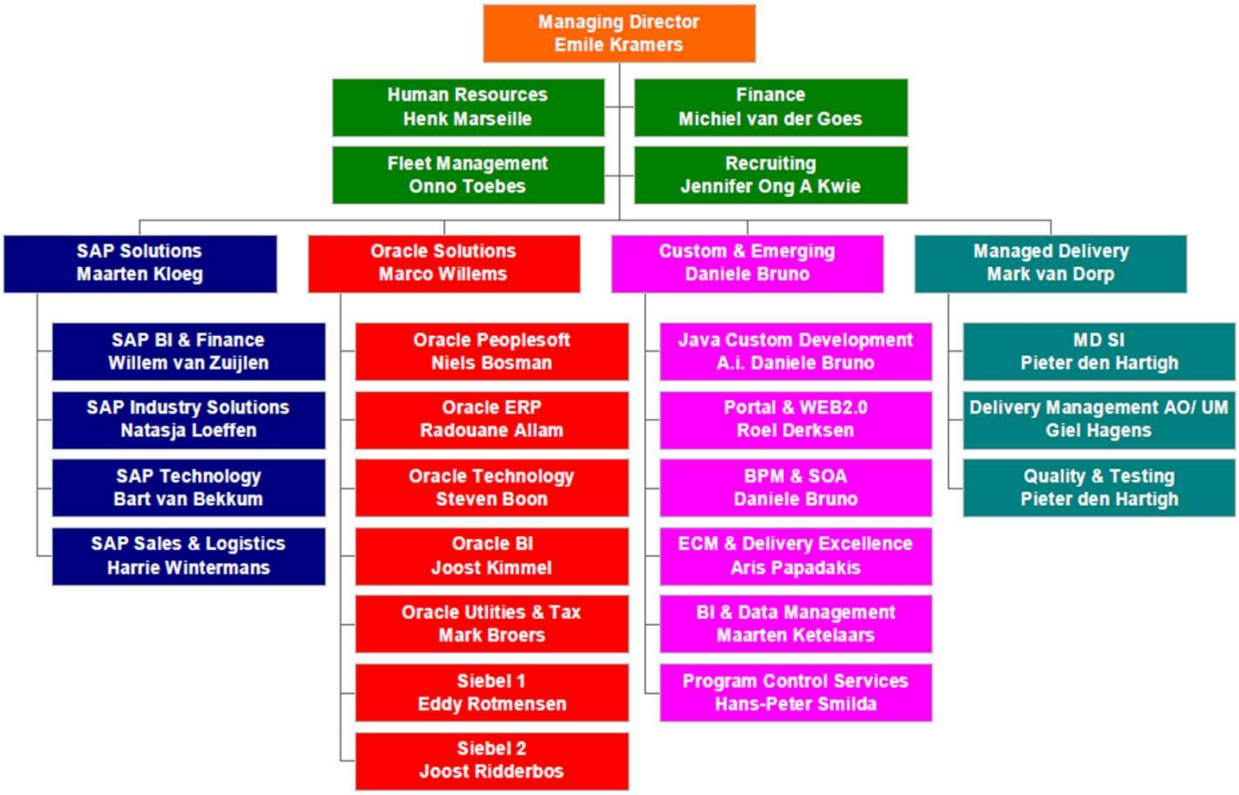
Afbeelding 19 - Organogram Accenture Technology Solutions (ATS)

Bijlage B. Vestigingen Accenture



Afbeelding 20 - Vestigingen Accenture

Bijlage C. Organisatieschema



Afbeelding 21 -Organisatieschema Accenture

Bijlage D. Welke analysetechniek in welke situatie

De keuze van de juiste analysetechniek is lastig te bepalen en afhankelijk van veel factoren. De keuze hangt af van de beschikbare data, het doel van de analyse en het ervaringsniveau van de data-analist.

Wanneer het de bedoeling is van de analyse om snel en globaal inzicht te verkrijgen in een analysebestand, dan verdient het aanbeveling om te beginnen met een aantal eenvoudige beschrijvende technieken. Dit kost relatie weinig tijd en is over het algemeen simpel uit te voeren.

Voor het voorspellen van de kans op het weglopen van klanten of de kans op productaanschaf is een voorspellende techniek vereist. Als tevens de interpretatie van het gebouwde model belangrijk wordt geacht, dan kan vaak het beste voor een regressie of een beslisboom worden gekozen. Als de interpretatie minder belangrijk is en simpelweg het beste model dient te worden gekozen, dan zijn in principe alle in het softwarepakket beschikbare voorspellende technieken geschikt om een model mee te bouwen, waarna de best voorspellende techniek uiteindelijk gebruikt wordt.

Bijlage E. Analysemethodes

Door middel van analysemethoden kunnen de verzamelde gegevens worden geanalyseerd. In dit hoofdstuk worden de analysemethodes: klant- en fraudeanalyse uitgewerkt.

1. Klant analyse

Een klantanalyse richt zich op het analyseren, toepassen en opslaan van kennis over klanten en manieren om klanten te benaderen. Veelal wordt hiervoor gebruik gemaakt van databases, statistische tools, datamining, Business Intelligence en rapportage methodieken.
De klant is tegenwoordig koning, waardoor focussen op een succesvolle één op één relatie noodzaak is geworden. Om de maximale klantwaarde te behalen is het van belang om te weten wat ze gaan doen, waar ze behoefte aan hebben en ze zodoende tevreden houden.
Het maximaliseren van de klantwaarde is een cruciaal punt geworden voor het maximaliseren van de algehele waarde van winstgevendheid van de organisatie. De beschikking over analytische tool is essentieel om de behoeften van de klant te voorspellen zodat de klanten tevreden worden gehouden en op hetzelfde moment de klantwaarde kan worden vergroot.

1.1 Doelstelling

De doelstelling van een klantanalyse is om:

- De klanten nog beter te kennen en smallere segmenten te bedienen;
- De juiste klanten te winnen op een kosten effectievere manier;
- Klantrelaties te verbeteren door te weten wat klanten willen;
- De juiste klanten langer te behouden.

1.1.1 Voorbeeld

Het voorbeeld uit tabel 15 is gebaseerd op een fictief bedrijf dat meer winstgevende resultaten in de toekomst wil bereiken door middel van de marketingcampagnes beter af te stemmen op de klant.
Het voorbeeld maakt gebruik van een beslissingslijst model om de kenmerken van klanten die het meest reageren te identificeren, gebaseerd op voorgaande acties.

Segment	Score	Cover...	Frequency	Probability
income income > 55.267	1	643	551	85,69%
number_transactions, rfm_score number_transactions > 2 and rfm_score > 12,33	1	533	206	38,65%
age age > 30 and age <= 34	1	505	111	21,98%
rfm_score rfm_score > 10,53	1	1.220	253	20,74%
Remainder		2.298	375	16,32%

Tabel 15 - Beslissingslijst klantanalyse

In de afbeelding is een deel van de uitkomst weergegeven. In de beslissingslijst staan verschillende segmenten, waarbij in elk segment een voorwaarde wordt weergegeven met het percentage reacties erachter. Onder het eerste segment vallen klanten met een inkomen van 55.267 of meer. Van deze klanten reageerden bij voorgaande campagnes 551 personen (Frequence) van de 643 die benaderd waren (Cover). Aan de hand van deze gegevens wordt het succes percentage berekend (Probability). Met behulp van deze tabel kan er worden bepaald welke klanten hoogstwaarschijnlijk zullen reageren en bij een volgende campagne sowieso benaderd moeten worden.

2. Fraude

Frauderen is niets anders dan een vorm van diefstal. Waar diefstal traditioneel geassocieerd wordt met het ontvreemden van materiële goederen, met name liquide geld, wordt fraude veelal in verband gebracht met misdadigheid in kringen waar veel geld in omgaat. Voorbeelden hiervan zijn het stelen van informatie met het oog op chantage, of het vervreemden van geld via illegale financiële transacties.

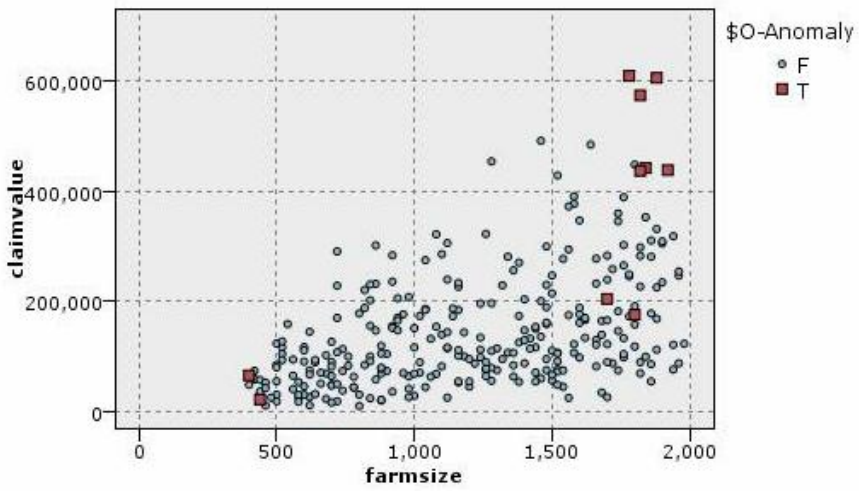
Banken, verzekeraars en overheidsinstanties signaleren zowel een toename in het aantal als in de geraffineerdheid van frauduleuze activiteiten. Om fraude effectief te bestrijden moeten organisaties continue hun processen verbeteren om klantgedrag in verschillende accounts en systemen te monitoren. Daarbij wordt onder meer gebruik gemaakt van voorspellende modellen.

2.1 Doelstelling

De doelstelling is om door middel van geavanceerde statistische en datamining technieken, fraude proberen te bestrijden en te voorkomen. Door deze technieken worden grote hoeveelheden data geanalyseerd om zo tijdig patronen en afwijkingen op te sporen. Hiermee worden voorspellende modellen ontwikkeld die vaststellen welke klanten of activiteiten mogelijk frauduleus zijn. De inzet van predictive analytics zorgt ervoor dat er automatisch voorspellingen en aanbevelingen worden gedaan waarop beleidsmakers tijdig kunnen anticiperen.

2.1.1 Voorbeeld

Het voorbeeld uit grafiek 6 heeft betrekking op het aanvragen van de landbouw subsidies. Het opsporen van gedragingen die kunnen wijzen op fraude worden in dit voorbeeld uitgewerkt door middel van anomalie (onregelmatigheden) detectie.



Grafiek 6 - Fraudeanalyse

In de grafiek is een anomalie detectie toegepast op de bedrijfsomvang tegenover de waarde van de claim. De rode gemarkeerde vierkantjes weergeven de afwijkende records oftewel de mogelijke fraudegevallen weer. Naast de hoge en lage claims die als mogelijke fraudegevallen worden weergeven, zijn er ook middenin de grafiek rode vierkantjes gemarkeerd. Deze onregelmatigheden zijn minder goed te plaatsen dan de andere en hebben waarschijnlijk geen betrekking op de variabele bedrijfsomvang maar te maken met een andere variabele.

Bijlage F. Modellering

Een model is een serie van regels, formules of vergelijkingen die kunnen worden gebruikt om een resultaat op basis van een reeks van input velden of variabelen te voorspellen.

In Predictive Workbench worden een verscheidenheid van modellering methodes aangeboden waarmee nieuwe informatie uit de gegevens kan worden gehaald en voorspellende modellen kunnen worden ontwikkeld. Elke methode heeft bepaalde sterke punten en is het meest geschikt voor bepaalde typen van problemen.

Voor het bouwen van een model zijn er minimaal drie elementen nodig:

- Een bron knooppunt die de gegevens van een externe bron inleest;
- Een type knooppunt dat veldeigenschappen zoals data type specificeert;
- Een modellering knooppunt dat een model genereert als de stroom van gegevens wordt uitgevoerd.

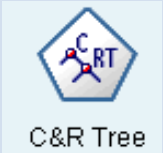
Daarnaast zijn er ook nog tabel en analyse knooppunten die gebruikt worden om een overzicht van de resultaten te weergeven nadat het model is gemaakt.

Modellering knooppunten zijn ingedeeld in classificatie, associatie en segmentatie modules zoals hieronder beschreven.

Classificatie module

De classificatie module helpt om een resultaat te voorspellen zoals de vraag of een klant blijft of weggaat of dat een transactie past bij een bekend patroon van fraude.

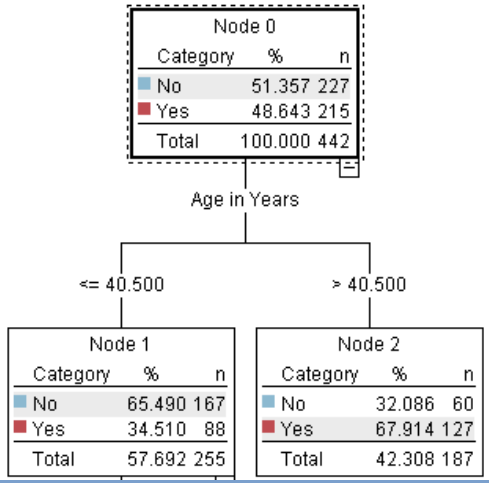
Classificatie en regressie boom



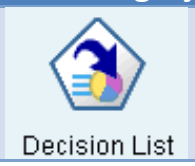
Het classificatie en regressie (C&R) boom knooppunt, is een beslissingsboom die het mogelijk maakt om voorspellingen te maken of toekomstige waarnemingen te classificeren.
De C&R boom knooppunt begint met het onderzoeken van de invoervelden om de beste verdeling te vinden. Deze verdeling definieert twee subgroepen, elk van deze subgroepen wordt vervolgens weer verdeeld in twee subgroepen en zo verder, totdat een van de stop criteria wordt geactiveerd. Alle verdelingen zijn binaire (slechts twee subgroepen).

Voorbeeld

Een telecom bedrijf wil graag kunnen bepalen welke klanten een abonnement op een interactieve nieuws dienst zouden kopen. Met behulp van de data kan er een voorspelling worden gemaakt aan de hand van de velden leeftijd, geslacht, opleiding, inkomen, uren telefoneren per week en het aantal kinderen. Door de toepassing van het C&R knooppunt kunnen de reacties worden geclassificeerd om de hoogste respons te bepalen voor de campagne.
In het voorbeeld hiernaast wordt er eerst een opsplitsing gemaakt op basis van leeftijd (jonger of gelijk aan 40,5 jaar (node 1) en ouder dan 40,5 jaar (node 2))



Beslissingslijst



De beslissingslijst knooppunt identificeert subgroepen of segmenten, die een hogere of lagere waarschijnlijkheid laten zien van een bepaalde binaire (ja of nee) uitkomst ten opzichte van de totale steekproef. Beslissingslijst modellen bestaan uit een lijst van regels waar elke regel een voorwaarde en een uitkomst heeft.

Voorbeeld

Het voorbeeld is gebaseerd op een fictief bedrijf dat meer winstgevende resultaten in de toekomst wil bereiken door middel van de marketingcampagnes beter af te stemmen op de klant.
Het voorbeeld maakt gebruik van een beslissingslijst model om de kenmerken van klanten die het meest reageren te identificeren, gebaseerd op voorgaande acties.

Dit model bevat een mix van hoge en lage presterende segmenten. Het model toont een succes percentage van 45,63% met een bereik van (1.577 hits van 3456 records).
In de beslissingslijst kan er per segment worden bekeken hoeveel personen er onder vallen (Cover (n)) en het aantal dat reageert (Frequency). Door middel van deze aantallen kan het succes percentage worden berekend (Probability). Aan de hand van deze gegevens kan er worden gekeken bij welke segmenten het succes percentage het hoogste is en dus de meeste kans heeft op een reactie (score van 1).

id	Segment	Score	Cover...	Frequency	Probability
	All segments including Remainder		13.504	1.952	14,45%
1	**months_customer** months_customer = "0"	Excluded	1.747	0	0,00%
2	**rfm_score** rfm_score <= 0	Excluded	6.003	0	0,00%
3	**rfm_score, income** rfm_score > 12,33 and income > 52.213	1	555	456	82,16%
4	**income** income > 55.267	1	643	551	85,69%
5	**number_transactions, rfm_score** number_transactions > 2 and rfm_score > 12,33	1	533	206	38,65%
6	**age** age > 30 and age <= 34	1	505	111	21,98%
7	**rfm_score** rfm_score > 10,53	1	1.220	253	20,74%
	Remainder		2.298	375	16,32%

Model Summary; Cover 3,456: Frequency 1,577: Probability 45.63%

Regressie

Het regressie knooppunt genereert een lineaire regressie model. Dit model maakt een schatting van de best passende lineaire vergelijking voor het voorspellen van de output, gebaseerd op de input velden. De regressievergelijking is een rechte lijn of vlak dat de gekwadrateerde verschillen tussen de voorspellende en daadwerkelijke output waarden minimaliseert. Dit is een veel voorkomende statistische techniek voor het samenvatten van gegevens en het maken van voorspellingen.

Logistische regressie is een statistische techniek voor de classificatie van records gebaseerd op waarden van invoervelden. Het is analoog aan lineaire regressie, maar neemt een doelveld met categorische waarden (kenmerken) aan in plaats van een numeriek bereik. Beide modellen, binomiale (voor doelen met twee categorieën) en multinomiale modellen (voor doelen met meer dan twee categorieën) worden ondersteund.
Logistische regressie werkt door middel van het opbouwen van een set van vergelijkingen die gerelateerd zijn aan elk van de uitvoerveld categorieën. Zodra het model is gegenereerd, kan het gebuikt worden om waarschijnlijkheden te voorspellen voor nieuwe gegevens.

Voorbeeld

In dit voorbeeld wordt het verlies van klanten (churn) voorspeld door gebruik te maken van de benodigde gegevens. Hiervoor wordt een binominaal model gebruikt omdat er 2 verschillende categorieën bestaan: klanten die blijven of klanten die gaan.

De classificatietabel hiernaast toont de resultaten voor het model. Bij elke stap wordt er een voorspeller toegevoegd. Na de eerste stap, wanneer er slecht één predictor is gebruikt, is de juistheid van de churn voorspelling van het model toegenomen van 0,0% naar 29,9%.

De laatste stap is stap 6. In dit stadium heeft het model besloten dat het niet langer nodig is om een predictor toe te voegen in het model. De nauwkeurigheid van voorspellingen voor degenen die weggaan als klant is gestegen van de oorspronkelijke 0% naar 47,1%. Dit is een belangrijke verbetering ten opzicht van de originele nul model dat geen predictors gebruikte.

	Observed		Predicted		
			Churn		% Correct
			No	Yes	
Step 1	**Churn**	**No**	668	58	92,0
		Yes	192	82	29,9
	Overall Percentage				**75,0**
Step 2	**Churn**	**No**	657	69	90,5
		Yes	160	114	41,6
	Overall Percentage				**77,1**
Step 3	**Churn**	**No**	661	65	91,0
		Yes	153	121	44,2
	Overall Percentage				**78,2**
Step 4	**Churn**	**No**	658	68	90,6
		Yes	148	126	46,0
	Overall Percentage				**78,4**
Step 5	**Churn**	**No**	658	68	90,6
		Yes	155	119	43,4
	Overall Percentage				**77,7**
Step 6	**Churn**	**No**	658	68	90,6
		Yes	145	129	47,1
	Overall Percentage				**78,7**

Neuraal netwerk

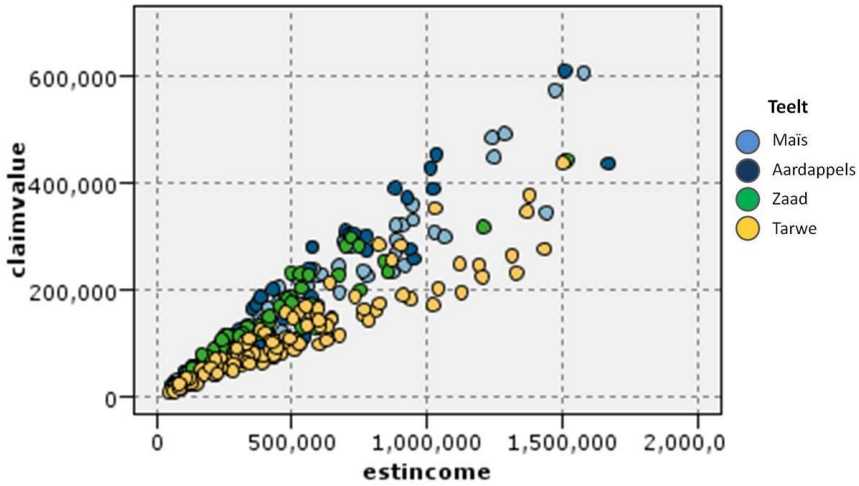
Een neuraal netwerk is in enige mate gebaseerd op de organisatie van het menselijke brein en de manier waarop de hersenen leren. Er zijn twee belangrijke structuurelementen in het neurale netwerk:

- Het knooppunt;
  - Te vergelijken met de zenuwcel in een menselijk brein;
- De koppeling;
  - Te vergelijken met de verbindingen tussen deze zenuwcellen.

Neurale Netwerken (NN) is een analyse methode die vooraf geen informatie nodig heeft over de verbanden tussen de input en output. De relaties tussen input variabelen en de output variabelen worden bepaald door de gegevens, en niet voorbepaald door een model.

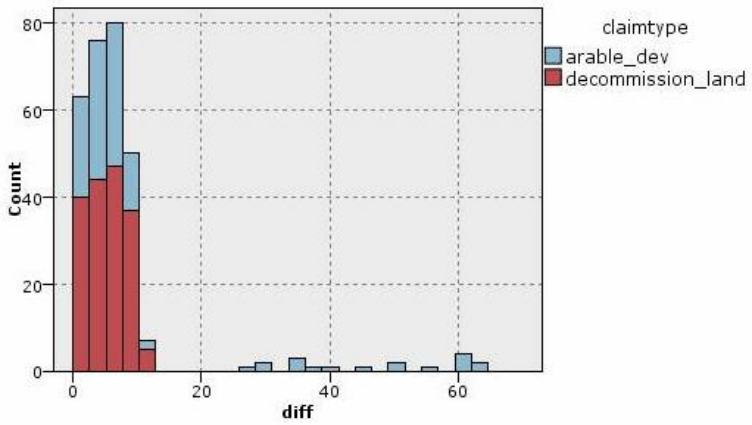
Voorbeeld

In het voorbeeld hiernaast wordt gescreend op mogelijke gevallen van fraude bij het aanvragen van subsidies voor de landbouw. Er wordt een neuraal netwerk gebruikt voor een verkenning van afwijkingen van de norm. Hierbij wordt er vooral gezocht naar subsidie aanvragen die veel (of te weinig) geld vorderen voor de omvang van het bedrijf.



In de grafiek hierboven is de voorspelde claim waarde uitgezet tegen de werkelijke waarde van de vordering.

Om het verschil tussen de werkelijke en de voorspelde waarde van claims te interpreteren wordt er gebruik gemaakt van een differentiatie model. Hierin is het verschil tussen de werkelijke waarde vergeleken met de voorspelde waarde van claims.



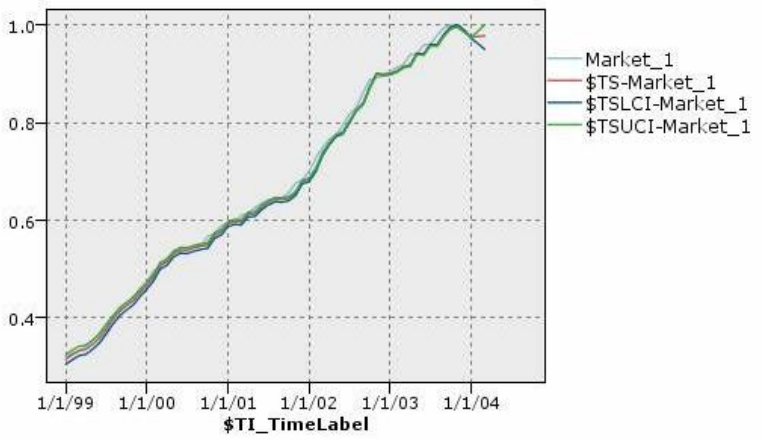
Time series

Het tijd series knooppunt maakt schattingen voor modellen van tijdreeksen en produceert prognoses op basis van tijdreeks gegevens. De volgende modellen worden hiervoor gebruikt:

- **Exponentiële groei**
  Gebruikt een gewogen gemiddelde van oude (vorige) en huidige waarden, waarbij het gewicht van de huidige waarden wordt aangepast om rekening te houden met schommelingen in de gegevens zoals seizoensinvloeden.
- **Autoregressief geïntegreerd bewegend gemiddelde (ARIMA)**
  De ARIMA tijdseries analyse gebruikt vertragingen en verschuivingen in historische gegevens om patronen (bijvoorbeeld seizoensgevoeligheid) te ontdekken en voorspellen.

Voorbeeld

In het voorbeeld wordt een prognose gemaakt van de abonnement gebruikers om het gebruik van bandbreedte te voorspellen. Prognoses worden gemaakt voor elk van de lokale markten die deel uitmaken van het nationale abonneebestand. Er wordt gebruik gemaakt van de tijdreeksmodellering. De prognose wordt uitgevoerd voor de komende drie maanden.



De lijn van $TS-Market_1 bevat de oorspronkelijke gegevens van het gegenereerde data model. $TSLCI-Market_1 bevat de waarden van een lagere betrouwbaarheidsinterval en $TSUCI-Market_1 de waarden van een hogere betrouwbaarheidsinterval.

De lijn voor de voorspelde en actuele gegevens over de gehele tijdreeks liggen erg dicht bij elkaar op de grafiek, waaraan je kunt zien dat dit een betrouwbaar model is.

Segmentatie

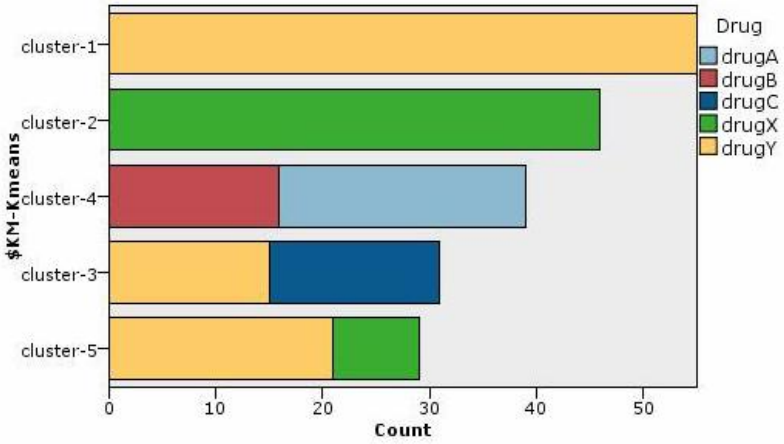
De segmentatie module wordt aanbevolen in gevallen waarin het specifieke resultaat nog onbekend is. Bijvoorbeeld bij de identificatie van nieuwe patronen van fraude, of wanneer het identificeren van groepen van belang is in een klantenbestand.

K-Means

Het K-Means knooppunt clustert de gegevens in verschillende groepen. De methode definieert een vast aantal clusters, wijst herhaaldelijk records aan clusters en stemt de clusters af totdat door middel van een verdere verfijning het model niet verbeterd kan worden. In plaats van een uitkomst proberen te voorspellen, gebruikt k-means een proces dat bekent staat om patronen te ontdekken in de set van invoervelden.

Voorbeeld

In het voorbeeld hiernaast zijn gegevens verzameld over een reeks patiënten die allemaal lijden aan dezelfde ziekte. Gedurende de behandeling reageert elke patiënt op één van de vijf medicijnen. Door middel van het K-Means model worden de gegevens geclusterd naar aanleiding van invoervelden. In de grafiek hiernaast zijn die clusters weergegeven met het medicijn dat de patiënten in dat cluster hebben gebruikt.



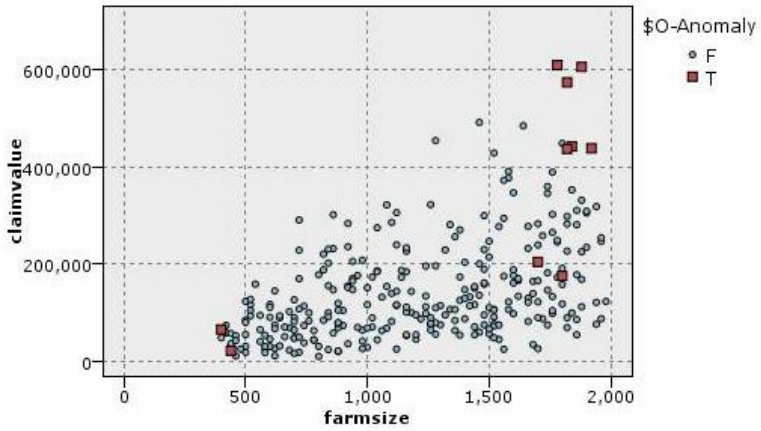
Anomalie detectie

Anomalie detectie modellen worden gebruikt om uitschieters of ongewone gevallen in de gegevens vast te stellen. In tegenstelling tot andere modellering methodes die ongebruikelijke regels opslaan, slaat het anomalie detectie model informatie op over hoe het normaal gesproken eruit ziet. Dit maakt het mogelijk om uitschieters te ontdekken zelfs als ze niet voldoen aan de bekende patronen. Dit kan nuttig zijn in toepassingen zoals het opsporen van fraude, waar nieuwe patronen continue worden ontdekt en gemaakt.

Anomalie Detectie kan een groot aantal velden onderzoeken om clusters te identificeren waarin soortgelijke records vallen. Elk record kan vervolgens worden vergeleken met anderen in de cluster om mogelijke afwijkingen te identificeren.

Voorbeeld

Het voorbeeld heeft betrekking op het aanvragen van de landbouw subsidies. Het opsporen van gedragingen die kunnen wijzen op fraude worden in dit voorbeeld uitgewerkt door middel van anomalie detectie.



In de grafiek hiernaast is een anomalie detectie toegepast op de bedrijfsomvang tegenover de waarde van de claim. De rode gemarkeerde vierkantjes weergeven de afwijkende records oftewel de mogelijke fraudegevallen weer. Naast de hoge en lage claims die als mogelijke fraudegevallen worden weergeven, zijn er ook middenin de grafiek rode vierkantjes gemarkeerd. Deze onregelmatigheden zijn minder goed te plaatsen dan de andere en hebben waarschijnlijk geen betrekking op de variabele bedrijfsomvang maar te maken met een andere variabele.

Associatie module

De Associatie module is vooral nuttig bij het voorspellen van meerdere uitkomsten. Bijvoorbeeld klanten die product X kochten, kochten ook product X en Z. Associatie modellen koppelen een bepaalde conclusie (zoals het besluit om iets te kopen) met een reeks voorwaarden.

Generaliseerde regel inductie

Het generaliseerde regel inductie (GRI) knooppunt ontdekt associatie regels in de data. GRI kan omgaan met zowel numerieke en categorische invoerwaarden.

Voorbeeld

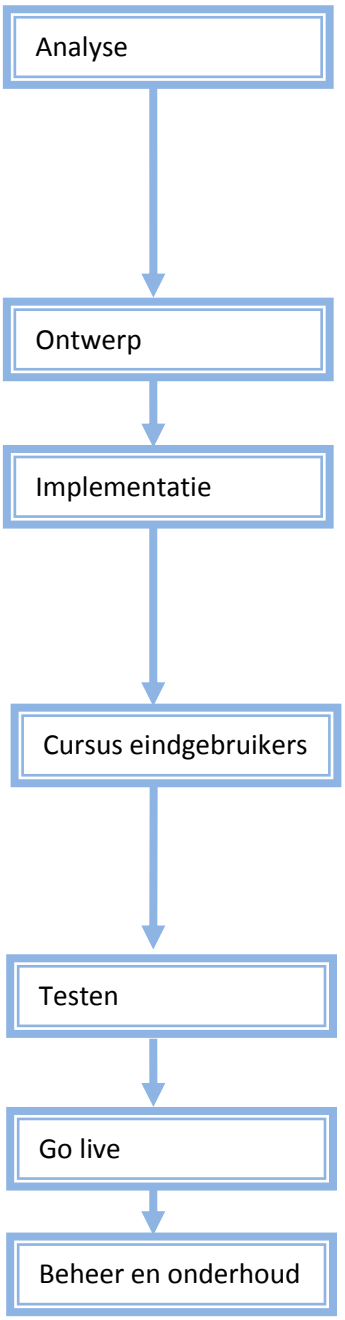
In het voorbeeld hiernaast wordt weergeven wat er wordt gekocht door klanten (Antecedent) en wat er daarnaast vaak bij wordt gekocht (Consequent).

- Bij support wordt het percentage weergegeven waarbij de Antecedenten kloppen;
- Bij Confidence wordt aangegeven bij hoeveel procent van de gevallen deze associatieregel klopt.

Consequent	Antecedent	Support %	Confidence %
frozenmeal	cannedveg beer	16.7	87.43
cannedveg	frozenmeal beer	17.0	85.88
beer	cannedveg frozenmeal	17.3	84.39
frozenmeal	beer	29.3	58.02
beer	frozenmeal	30.2	56.29

Bijlage G. Implementatie

De fasen die worden gevolgd voor de implementatie bestaan uit:



In deze fase gaat het project van start met een onderzoek naar de mogelijkheden in het bedrijf. Hierbij is het van belang om te weten welke systemen er al aanwezig zijn in de organisatie, hoe de processen in de organisatie lopen en wat het doel is van de implementatie. Door middel van oriënterende gesprekken met de klant worden de requirements (wensen) kenbaar gemaakt. Hierop volgend wordt het projectmanagementplan opgesteld waarin de loop van het project wordt weergegeven.

Op basis van de goedgekeurde specificaties zal het project worden ontworpen. Hieronder worden zowel technische delen van het project verstaan alsmede de onderdelen die voor de eindgebruiker belangrijk zijn.

Het ontwerp wordt vervolgens in een logische volgorde volgens de documentatie uit de ontwerpfase geïmplementeerd. In deze fase wordt bijvoorbeeld de koppeling van de systemen gemaakt en de tool wordt ingericht met de data uit de systemen. Als het wenselijk is kan deze fase opgedeeld worden in een aantal kleinere fasen waarin de opdrachtgever wat frequenter het projectverloop kan zien.

Deze cursus biedt een overzicht van datamining en de grondbeginsel van het gebruik van Predictive Workbench. Met behulp van de CRISP-DM-methodologie (paragraaf 7.3.2) worden de principes en de praktijk van datamining toegelicht. Deze cursus demonstreert de manier waarop modellen voor de voorspellingen worden ontwikkeld waardoor de eindgebruikers er na de implementatie zelf mee kunnen werken.

In deze fase wordt er door zowel de testers als de klant gecontroleerd of alle ontwerpen correct zijn gebouwd volgens de afspraak. Ook kunnen er in deze fase fouten boven water komen die al in eerdere stadia gemaakt zijn.

Het systeem is klaar en getest. Het bedrijf kan de tool in gebruik nemen.

De klant heeft de keuze om het onderhoud zelf te verrichten of dit eventueel om een bepaalde periode uit te laten voeren door Accenture.

Bijlage H. Verwachte voordelen/nadelen

SAP BO Predictive Workbench helpt organisaties om beter inzicht te krijgen in de data om vervolgens de churn beter te kunnen voorspellen.

Resultaten kunnen worden gemodelleerd zodat kan worden bekeken welke factoren invloed hebben op de resultaten zodat er proactief kan worden gereageerd en kansen kunnen worden benut om risico's te beperken. Dit kan vervolgens leiden tot:

- Eenvoudigere en snellere klantservice processen;
- Hogere klanttevredenheid;
- Hogere conversie van campagnes of folders;
- Beter inzicht in klantbehoeften;
- Betere segmentatie van klanten.

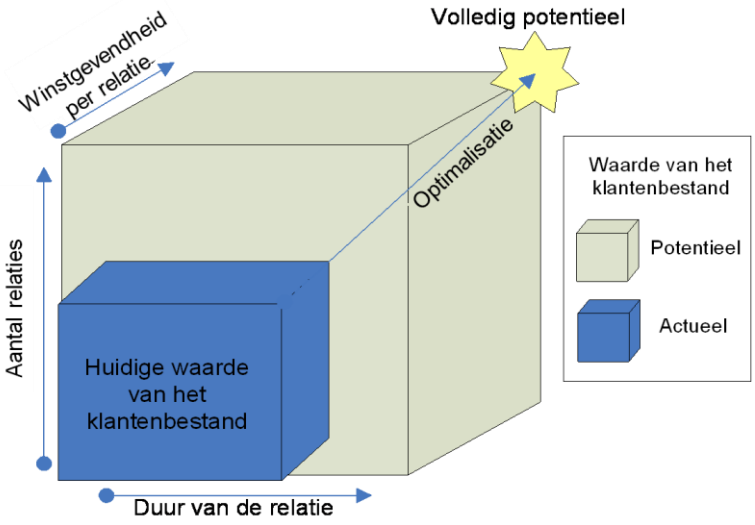
Deze business case is gericht op het voorspellen van churn maar hiernaast is er nog veel meer mogelijk met de tool die een voordeel kunnen opleveren voor het bedrijf.
Het uiteindelijke doel van de tool is dat het bedrijf zich kan ontwikkelen tot een 'Predictive Enterprise': Een organisatie die beslissingen stuurt en optimaliseert door het integreren van voorspellende technologie in de bestaande operatie met als doel de winstgevendheid te maximaliseren.
Het grootste verwachte nadeel is de hoge licentiekosten die voor de tool moeten worden betaald.

Bijlage I. Kosten[16]

In dit hoofdstuk is een voorbeeld uitgewerkt waarin de klantwaarde wordt berekend aan de hand van de omzet en kosten die een bedrijf maakt. Eerst wordt er ingegaan op de definitie van de klantwaarde in het bedrijf en vervolgens op de winstgevendheid van een gemiddelde klant.

1. Klantwaarde

Klantwaarde kan gedefinieerd worden als de netto contante winst die een klant oplevert over een bepaalde periode. De totale waarde van een klantenbestand is de optelling van de individuele klantwaarden of waarden van bepaalde klantgroepen. Om de factoren die de waarden van het klantbestand beïnvloeden, te begrijpen wordt er gebruikt gemaakt van een klantwaarde kubus (afbeelding 22). In deze grafische weergaven is te zien dat de mogelijke waarde van het klantenbestand (volledig potentieel) de einduitkomst is van het maximale aantal relaties (klanten), de maximale winstgevendheid per relatie en de maximale duur van de relatie.



Afbeelding 22 - Kubus klantwaarde

Dit betekent dat er in principe drie manieren zijn om de waarde van het klantenbestand te maximaliseren: acquisitie, ontwikkeling en retentie.

- Acquisitie is het werven van nieuwe klanten. Dit is de dimensie van de kubus waar de meeste bedrijven zich vroeger voornamelijk op richten, onvoldoende rekening houdend met het feit dat klanten ook verder ontwikkelen en vastgehouden dienen te worden;
- Ontwikkeling betreft het optimaliseren van de winstbijdrage van de klant;
- Retentie betreft zoals al eerder vermeld in de business case het vasthouden van de klant, oftewel het verlengen van de relatie met de klant.

[16] Literatuurlijst: bron 3

2. Berekening winstgevendheid

Voor de berekeningsmethodiek die hieronder is uitgewerkt is het Lifetime Value (klantwaarde) model gebruikt. Dit model berekent de langetermijneffecten en houdt daarbij rekening met alle belangrijke factoren zoals acquisitie, ontwikkeling, retentie, variabele kosten, vaste kosten (o.a. kosten tool).

De klantwaarde kan berekend worden met behulp van enkele factoren:

Retentieratio — Het aantal klanten dat zowel in de huidige als voorgaande periodes kochten, gedeeld door het aantal klanten in de voorgaande periode.

Gemiddelde bestedingen per klant — Dit kan berekend worden door de totale omzet te delen door het aantal klanten in een jaar.

Variabele kosten — Alle kosten die stijgen bij het toenemen van het aantal klanten, bijvoorbeeld de kostprijs van producten en/of diensten, maar ook de kosten van klantenservice. Variabele kosten kunnen berekend worden als een bedrag of een percentage van de gemiddelde bestedingen.

Acquisitiekosten — Kosten die gemaakt worden om een nieuwe klant te 'winnen'. Dat zouden bijvoorbeeld de kosten kunnen zijn van een wervende marketingcampagne inclusief welkomstaanbod of de kosten van de buitendienst die klanten werft.

Netto contante waarde en discontoratio — Geld dat er in de toekomst mogelijk verdiend wordt, is niet zo veel waard als geld dat momenteel in bezit is. Dit is een economisch principe. Dergelijke toekomstige opbrengsten moeten daarom worden 'gekort'. Door de toekomstige waarde te delen door een discontoratio wordt de netto contante waarde verkregen.

Aan deze factoren zitten wel nog een aantal 'wetmatigheden' vast:

- De retentieratio stijgt door de jaren heen. Dit komt doordat klanten naarmate ze langer bij het bedrijf goederen en/of diensten kopen in theorie steeds trouwer worden;
- Naarmate men langer klant is gaat men vaak meer besteden; men koop meer producten of duurdere producten, resulterend in een toename van de gemiddelde besteding door de jaren heen;
- De variabele kosten nemen vaak af naarmate men langer klant is.

De berekening uit tabel 17 is gemaakt aan de hand van de variabelen uit tabel 16:

Variabelen	Waarden	Variabelen	Waarden	Klanten
Initieel aantal klanten	21.422	Acquisitiekosten per klant	€ 25,00	
Retentiegraad	70%	Retentiekosten per klant (groep 1)	€ 7,50	26,9%
Groeifactor retentiegraad	3,0%	Retentiekosten per klant (groep 2)	€ 5,00	3,3%
Jaarlijks aantal klanten	4.500	Retentiekosten per klant (groep 3)	€ 2,50	39,9%
Disconterings percentage	8%	Vaste kosten (jaarlijks)	200.000	
Gemiddelde klantbesteding per maand	€ 30,00	Groeifactor vaste kosten	1%	
Variabele kosten (percentage omzet)	20%			

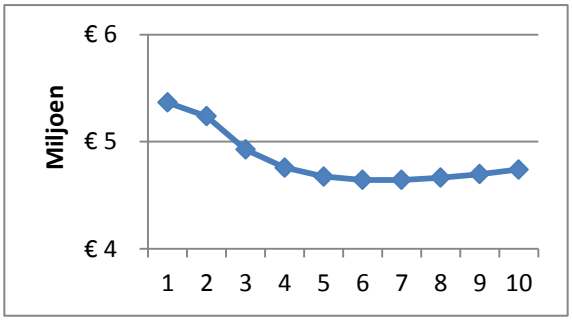
Tabel 16 - Variabelen berekening winst

Jaar	1	2	3	4	5
Aantal klanten	21.422	14.995	10.812	8.029	6.142
Retentieratio	*70%*	*72%*	*74%*	*76%*	*79%*
Nieuwe klanten		4.500	3.150	2.271	1.687
		70%	72%	74%	76%
			4.500	3.150	2.271
			70%	72%	74%
				4.500	3.150
				70%	72%
					4.500
					70%
Totaal aantal klanten	**21.422**	**19.495**	**18.462**	**17.950**	**17.749**
Gemiddelde besteding klant	€ 360	€ 360	€ 360	€ 360	€ 360
Omzet	**€ 7.711.920**	**€ 7.018.344**	**€ 6.646.206**	**€ 6.462.083**	**€ 6.389.744**
Variabele kosten	€ 1.542.384	€ 1.403.669	€ 1.329.241	€ 1.292.417	€ 1.277.949
Acquisitiekosten	€ 535.550	€ 112.500	€ 112.500	€ 112.500	€ 112.500
Retentiekosten groep 1	€ 43.219	€ 39.332	€ 37.246	€ 36.215	€ 35.809
Retentiekosten groep 2	€ 3.535	€ 3.217	€ 3.046	€ 2.962	€ 2.929
Retentiekosten groep 3	€ 21.368	€ 19.447	€ 18.416	€ 17.905	€ 17.705
Vaste kosten	€ 200.000	€ 202.000	€ 204.020	€ 206.060	€ 208.121
Winst	**€ 5.365.864**	**€ 5.238.180**	**€ 4.941.737**	**€ 4.794.024**	**€ 4.734.732**
Contante waarde bruto marge	€ 5.163.305	€ 4.667.075	€ 4.076.808	€ 3.661.990	€ 3.348.794
Klantwaarde	€ 241	€ 239	€ 221	€ 204	€ 189

Tabel 17 - Berekening winst

2.1 Toelichting

Het totaal aantal nieuw geworven klanten in jaar 1 bestaat uit 21.422 klanten.
De gemiddelde besteding per klant is € 30,- per maand dus € 360,- per jaar. In dit geval blijft de gemiddelde besteding door de jaren heen gelijk.
De omzet per jaar kan berekend worden door het totaal aantal klanten in dat jaar te vermenigvuldigen met € 360,-.
De variabele kosten worden berekend door de omzet te vermenigvuldigen met 20%.
De acquisitiekosten zijn per nieuwe klant €25,-. Daarnaast zijn er nog retentiekosten en vaste kosten waar in de volgende paragraaf op ingegaan wordt.
De winst in een jaar wordt vervolgens berekend door de kosten van de omzet af te trekken (grafiek 7).

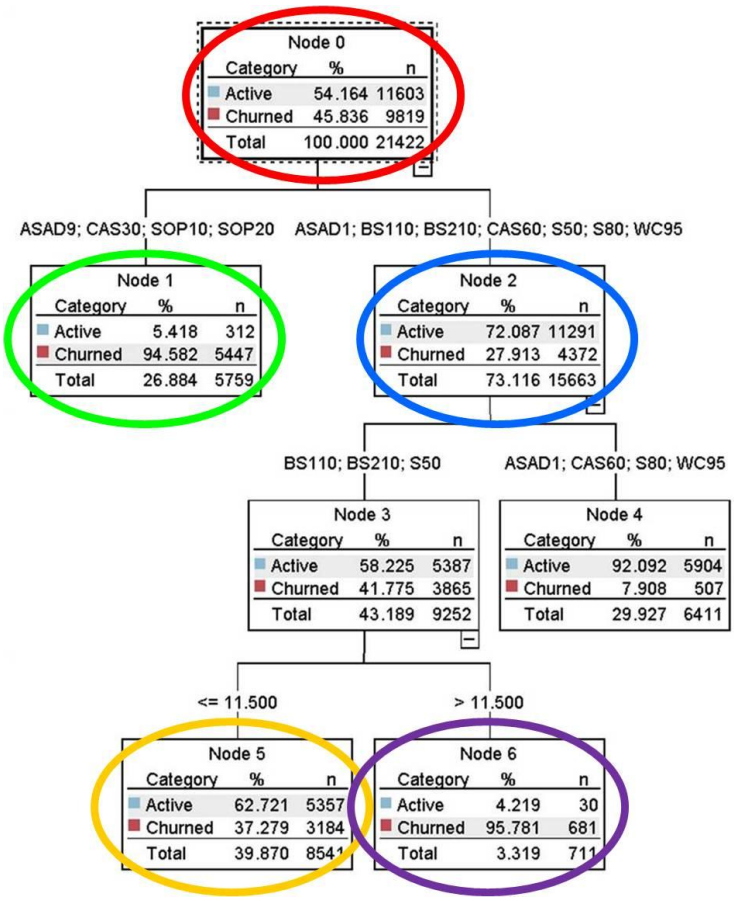


Grafiek 7 - Winst per jaar

Voor het gemak wordt ervan uitgegaan dat de rente plus de risicofactor 10% bedraagt. Vervolgens kan de netto contante waarde van de winst per jaar berekend worden door de winst te delen door het disconterings percentage.

Retentiekosten

De retentiekosten zijn opgedeeld in 3 categorieën. Deze categorieën zijn ontstaan naar aanleiding van de beslisboom (afbeelding 23) uit paragraaf 10.2.1. In node 0 (rode cirkel) is de kans op churn weergegeven van alle klanten. Vervolgens wordt er een opsplitsing gemaakt aan de hand van verschillende abonnementen. Wanneer klanten een abonnement hebben die bij node 1 (groene cirkel) is geselecteerd, is de kans op churn 95% en bij node 2 (blauwe cirkel) is de kans op churn slechts 28%. Omdat de kans op churn zo hoog is bij node 1 wordt de beslisboom hieronder niet verder opgesplitst. Dit in tegenstelling tot node 2 waar aan de hand van andere input velden de kans op churn nog verder wordt uitgediept. Hierdoor kan er aan de hand van andere eigenschappen van klanten worden gekeken hoe groot de kans op churn is.



Afbeelding 23 - Beslisboom kans op churn

Bij node 1 (groene cirkel) en 6 (paarse cirkel) is de kans op churn heel erg hoog, (meer dan 94%) waardoor er voor deze groep een gerichte retentie actie wordt ingezet met als doel om de klanten te behouden. Bij de klanten die onder node 5 (gele cirkel) vallen is de kans op churn 37%. In vergelijking met node 1 en 6 lijkt dit weinig maar ook hierbij is de kans op churn redelijk aanwezig, waardoor er ook hiervoor een (kleinere) actie ingezet kan worden.

Doordat onder node 1 de meeste klanten vallen en de kans op churn zo hoog is, wordt hiervoor een retentie campagne opgezet waarvan de kosten €7,50 zijn per klant.
Bij node 6 is de kans ook hoog maar is het percentage klanten dat hieronder valt wat kleiner waardoor er hier een campagne wordt ingezet die slechts €5,00 kost per klant.
En als laatste node 5 waarbij de kans op churn gering is maar wel echt aanwezig, wordt er een campagne ingezet die €2,50 kost per klant.

Hierdoor kan de volgende verdeling van retentiekosten worden gemaakt (tabel 18).

Node	Aantal klanten	Percentage churn	Kosten per klant
1	5.759	95%	€ 7,50
5	8.541	37%	€ 2,50
6	3.319	96%	€ 5,00

Tabel 18 - Verdeling retentiekosten

Deze kosten worden vervolgens doorberekend aan de hand van het aantal klanten die hieronder vallen en dit bedrag wordt per jaar opgenomen in de totale berekening.

Vaste kosten

De vaste kosten bestaan uit het totaalbedrag per jaar dat ongevoelig is voor veranderingen in de omvang van de bedrijfsactiviteiten. In de vaste kosten worden onder andere de kosten van de totale implementatie van de tool Predictive Workbench opgenomen (tabel 19).

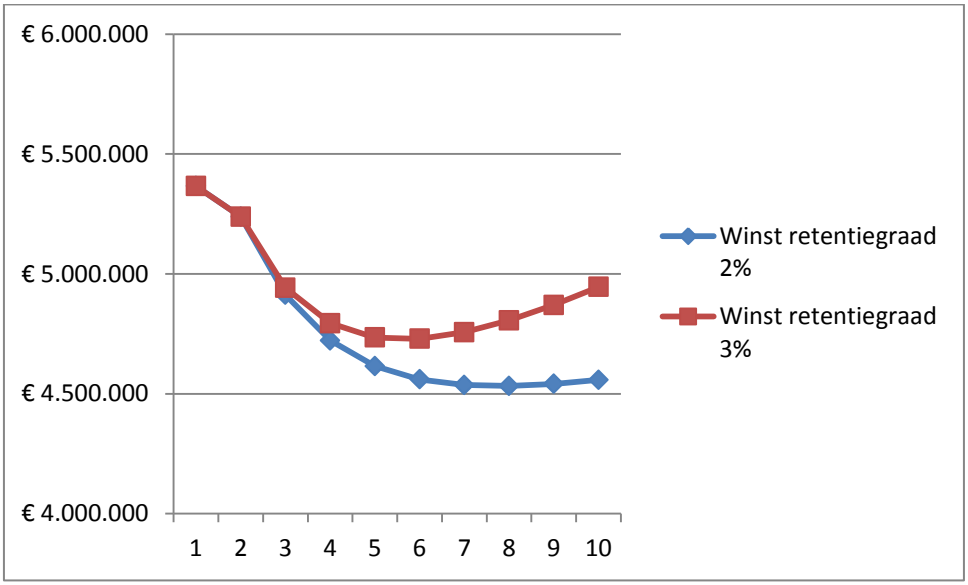
Beschrijving	Berekening		Kosten
Cursus	€ 550,- * 5 dagen = € 2.750,- dit bedrag * 2		€ 5.500,-
Licentie tool	Bedrag gebaseerd op 2 medewerkers van het bedrijf per jaar		€ 37.500,-
Onderhoud	15% * kosten licentie per jaar		€ 5.625,-
Totaal		**Eenmalig**	**€ 5.500,-**
		Jaarlijks	**€ 43.125,-**

Tabel 19 - Kosten tool

3. Investeringsberekening

De berekening van de investering wordt gemaakt aan de hand van de contante waarde van de onderneming. In dit voorbeeld wordt hieronder de totale winst verstaan van een bedrijf over een tijdsperiode van 5 jaar (de waarde van een bedrijf) rekening houdend met een disconteringvoet (rendementspercentage). Deze wordt meegenomen in de berekening omdat geld dat in de toekomst verdiend wordt op dit moment minder waard is dan geld een onderneming op dit moment in bezit heeft.

In grafiek 8 is het verschil in winst zichtbaar wanneer de retentiegraad met 1 procent wordt verhoogd. In deze grafiek wordt het belang van retentie duidelijk. Door middel van de tool Predictive Workbench is het gemakkelijk om deze specifieke groep te benaderen met een aantrekkelijke aanbieding in de hoop dat klanten langer blijven.



Grafiek 8 - Winst verschillende retentiegraden

Met de verhoging van de retentiegraad gaat de cumulatieve contante waarde van de onderneming van €20.753.538,- naar €20.917.971,- wat een verschil is van €164.433,-. Dit verschil wordt verkregen door de mogelijkheden van de tool SAP BO Predictie Workbench toe te passen op het probleem van churn in het bedrijf.

In deze business case is er nog maar één mogelijkheid van de tool uitgewerkt, er zijn veel meer mogelijkheden om ook andere problemen op te lossen. Doordat er alleen beschikking was over demogegevens van een telecom bedrijf, is ervoor gekozen om alleen dit onderwerp gedetailleerd uit te werken.

Bijlage J. Demo

Voor het bouwen van een model in SAP BO Predictive Workbench zijn er minimaal drie elementen nodig:

- Een bron knooppunt die de gegevens van een externe bron inleest;
- Een type knooppunt dat veldeigenschappen zoals data type definieert;
- Een modellering knooppunt dat een model genereert als de gegevensstroom wordt uitgevoerd.

Daarnaast zijn er ook nog tabel en analyse knooppunten die gebruikt worden om een overzicht van de resultaten te weergeven nadat het model is gemaakt.
In afbeelding 24 is weergeven hoe de gegevens van het telecom bedrijf in de tool zijn opgebouwd.

Brongegevens

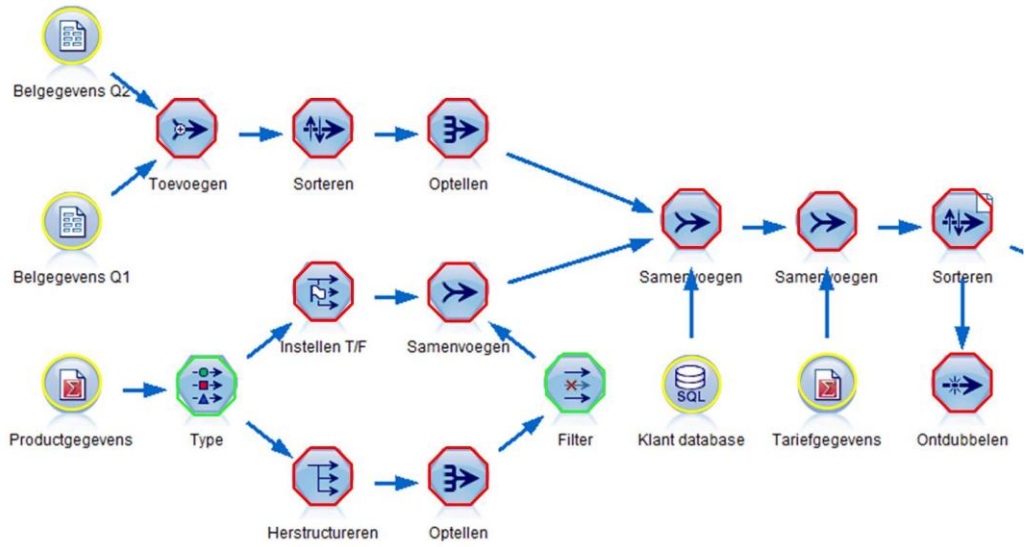
De noden met gele rondjes zijn de bronnen waar gegevens van het telecom bedrijf zijn ingeladen. De belgegevens zijn toegevoegd door middel van een variabel bestand, de product- en tarief gegevens via een statistiek bestand en de klantgegevens door middel van een Acces database.

Record

De rode node hebben betrekking op de records. Aan de hand van deze bewerkingen kunnen gegevens worden toegevoegd, gesorteerd, opgeteld, samengevoegd. In dit voorbeeld is dit nodig in verband met de verschillende ingeladen bronnen met gegevens. Door middel van de bewerkingen wordt er één totaal overzicht gemaakt met gegevens waar vervolgens analyses van kunnen worden gemaakt.

Veld

De node met een groen randje stellen eigenschappen van de velden zelf in. Hierdoor kan er per veld worden gekeken of de meetwaarde nominaal, rangschikkend, continue of standaard zijn. En bij de filter worden waarden eruit gefilterd die niet van toepassing zijn op dit voorbeeld.

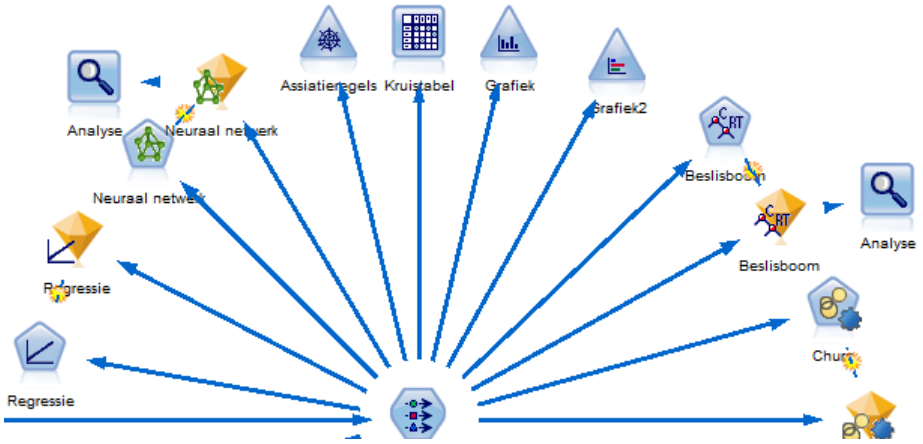


Afbeelding 24 - Gegevens telecom bedrijf

Modelleren

Een model is een serie van regels, formules of vergelijkingen die kunnen worden gebruikt om een resultaat op basis van een reeks van input velden of variabelen te voorspellen.

In Predictive Workbench worden een verscheidenheid van modellering methodes aangeboden waarmee nieuwe informatie uit de gegevens kan worden gehaald en voorspellende modellen kunnen worden ontwikkeld. Elke methode heeft bepaalde sterke punten en is het meest geschikt voor bepaalde typen van problemen.



Afbeelding 25 - Analysemodellen

Beschrijvende analyses

Kruistabel

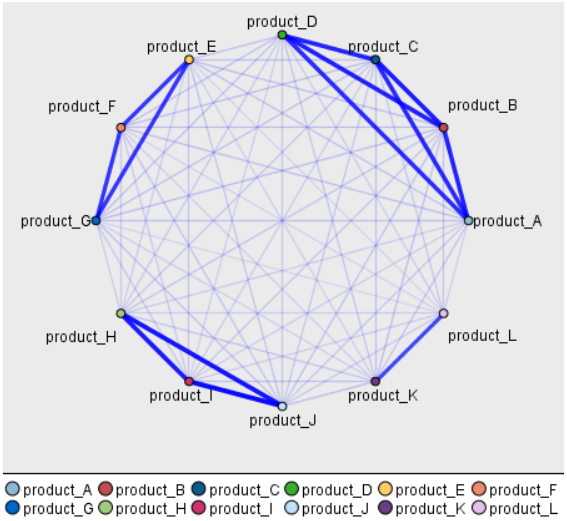
In tabel 20 (paragraaf 5.2.1) is het geslacht van klanten uitgezet tegenover de churn. In dit voorbeeld is weergeven dat van alle vrouwen de kans op churn 44% is en percentage van degene die blijven 56% is. Bij mannen ligt dit percentage op 52% kans op churn en dus 48% die blijven. In totaal (mannen en vrouwen) blijven er 54% en gaat 46% weg.

		Churn		
Gender		Active	Churned	Total
F	Count	8544	6776	15320
	Row %	55.770	44.230	100
M	Count	7967	7310	15277
	Row %	52.150	47.850	100
Total	Count	16511	14086	30597
	Row %	53.963	46.037	100

Tabel 20 – Kruistabel - churn/geslacht

Associatieregels

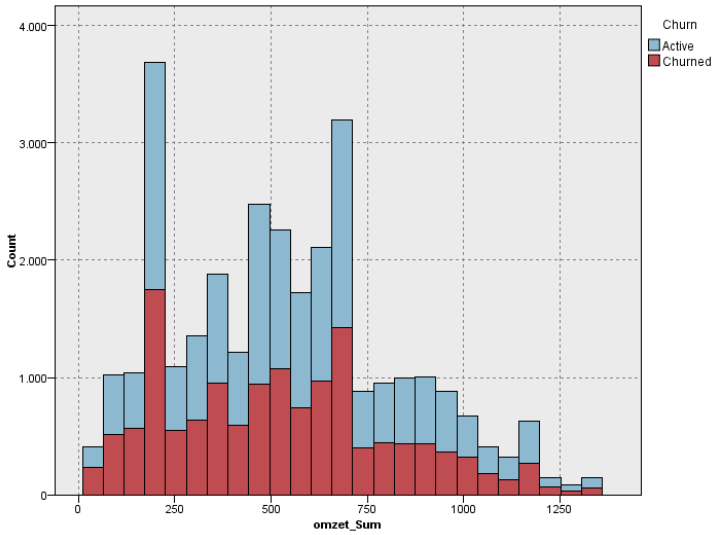
In grafiek 9 worden de associatieregels (paragraaf 5.2.3) weergegeven van producten. In deze grafiek is gekeken welke producten vaak in combinatie worden gekocht. Hieruit is af te leiden dat bijvoorbeeld producten E, F en G vaak samen worden aangeschaft. Door deze analyse kan er bijvoorbeeld worden gedacht over een combinatie pakket van deze producten, waardoor klanten bij de aanschaf van dit pakket een bepaalde korting krijgen. Dit is waarschijnlijk een interessant aanbod voor klanten waardoor de churn mogelijk kan worden verlaagd.



Grafiek 9 – Associatieregels - Producten

Grafieken

In deze paragraaf zijn een paar grafieken (paragraaf 5.2.2) uitgewerkt waardoor er in één oogopslag de mogelijke churn kan worden weergegeven.



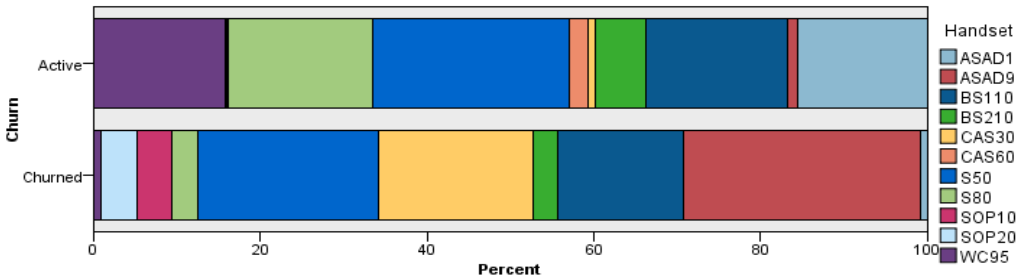
Grafiek 10 - Omzet

Omzet

In grafiek 10 is de totale omzet uitgezet tegenover de churn (rood). Hierdoor kan er worden gekeken hoe groot de kans op churn is aan de hand van de omzet die er behaald wordt bij klanten. In deze grafiek is te zien dat dit aantal bijna overal gelijk is. Alleen bij een omzet van net iets minder dan 500 is het aantal klanten dat blijft een stukje hoger dan de klanten die gaan.

Abonnement

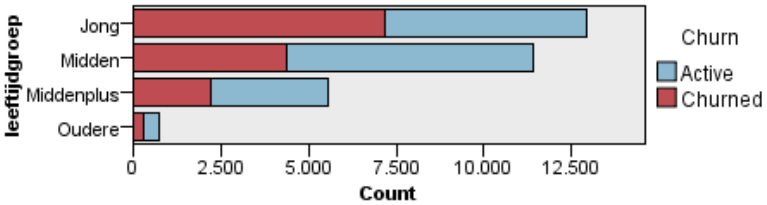
In grafiek 11 is de kans op churn weergeven op basis van het abonnement van klanten. Hierdoor is een snel overzicht te verkrijgen van de abonnementen waarbij vervolgens meteen kan worden gekeken of klanten bij het abonnement een grotere kans hebben om te blijven of te vertrekken. Bijvoorbeeld bij de klanten die de abonnementen ADAD1, ASAD9, CAS30 en SOP10 hebben, is het risico op churn een stuk hoger.



Grafiek 11 - Abonnement

Leeftijdsgroep

In grafiek 12 zijn de klanten opgedeeld in bepaalde leeftijdsgroepen. In dit voorbeeld is onderscheid gemaakt tussen:
Jong (<=25), Midden (<=40), Middenplus (<=60) en Oudere (>60). Van elke groep is weergegeven hoe groot de kans op churn is. Aan de hand van deze grafiek is te zien dat bij jongere de kans op churn het grootste is.



Grafiek 12 - Leeftijdsgroep

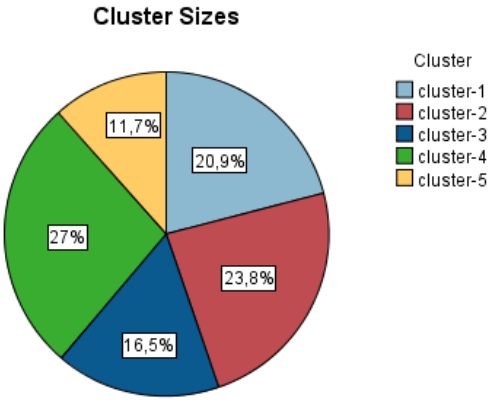
Cluster

Wanneer er een clusteranalyse (paragraaf 5.2.4) wordt uitgevoerd, kan dit in eerste instantie worden gedaan door middel van auto cluster. Op deze manier kan er worden gekeken welke clustermethode het beste resultaat weergeeft voor de geselecteerde input gegevens. Wanneer dit model is uitgevoerd, wordt de clustermethode K-means als beste model uitgekozen, welke in als eerste regel wordt weergeven in het overzicht (afbeelding 26).

Use?	Graph	Model	Build Time	...	Number of	Smallest Cluster (N)	Smallest Cluster (%)	Largest Cluster (N)	Largest Cluster (%)	Smallest/ Largest	Importance
☑		K-means 1	5	...	5	3288	10	8015	26	0,41	0
☐		TwoStep 1	5	...	5	3288	10	8441	27	0,39	0
☐		Kohonen 1	5	...	28	10	0	3288	10	0,003	0

Afbeelding 26 - Mogelijke clustermodellen

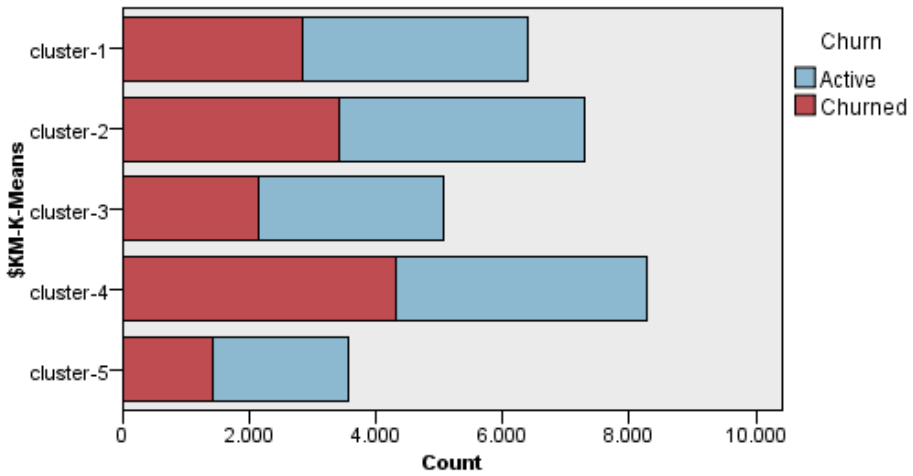
K-means is een clustermodel die patronen ontdekt in de set van invoervelden. De methode definieert clusters totdat het model niet verbeterd kan worden.
Na de uitvoering van het model wordt de data opgedeeld in vijf clusters die een bepaalde gelijkenis vertonen.



Grafiek 13 - Clusteranalyse

Het model geeft wel aan dat de kwaliteit van de clusteranalyse niet heel goed is (afbeelding 27). Er kan dus beter worden gekozen voor een ander model wat meer zekerheid biedt.

Afbeelding 27 - Kwaliteit clustermodel



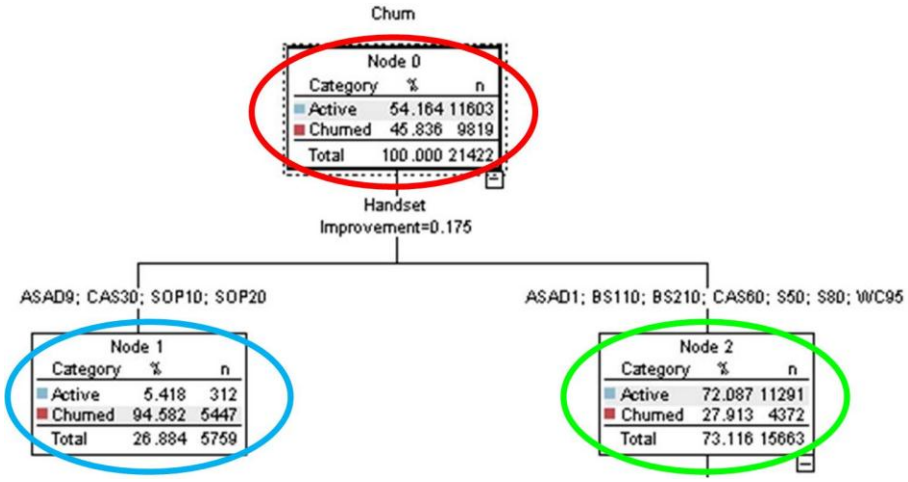
Grafiek 14 - Clusteranalyse churn

Om de churn in de clusters te achterhalen kan er een grafiek worden toegevoegd waar de clusters worden uitgezet tegenover de churn bij klanten (grafiek 13). Hierin is te zien dat de churn bij alle vijf de clusters ongeveer gelijk is waardoor aan de hand van deze analyse de churn nog niet nauwkeurig voorspeld kan worden.

Voorspellende analyses

Beslisboom

De beslisboom (paragraaf 5.3.1) in afbeelding 28 is opgesteld aan de hand van de inputwaarden tarief, geslacht, leeftijd en abonnement. Het doel van de beslisboom is het voorspellen van de churn. In de afbeelding wordt eerst node 0 (rode cirkel) weergegeven waarin de kans op churn staat weergegeven van alle klanten. Vervolgens wordt er een opsplitsing gemaakt aan de hand van de verschillende abonnementen. Wanneer klanten een abonnement hebben die bij node 1 (blauwe cirkel) is geselecteerd, is de kans op churn 95%. En bij node 2 is de kans op churn slechts 28% wanneer klanten een abonnement hebben die hiervoor geselecteerd zijn. Omdat de kans op churn zo hoog is bij node 1 wordt de beslisboom hieronder niet verder opgesplitst. Dit in tegenstelling tot node 2 waar aan de hand van andere input velden de kans op churn nog verder wordt uitgediept. Hierdoor kan er aan de hand van eigenschappen van klanten worden gekeken hoe groot de kans op churn is.



Afbeelding 28 - Beslisboom churn

Wanneer een analyse van de beslisboom wordt gemaakt, kan het resultaat van de juistheid van het model worden weergegeven (afbeelding 29). De analyse in dit voorbeeld laat zien dat 87% correct is voorspeld. Aan de hand van dit percentage is vast te stellen dat de betrouwbaarheid van dit model hoog is.

Results for output field Churn

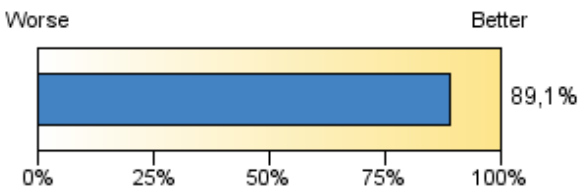
Comparing $R-Churn with Churn

Correct	26.653	87,11%
Wrong	3.944	12,89%
Total	30.597	

Afbeelding 29 - Resultaat churn

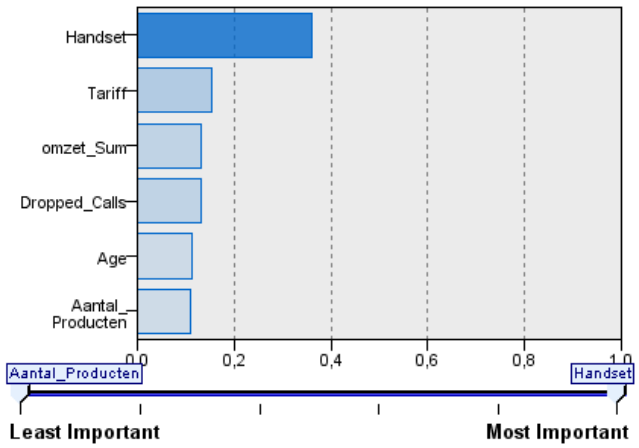
Neuraal netwerk

Neurale Netwerken (paragraaf 5.3.4) is een analyse methode die vooraf geen informatie nodig heeft over de verbanden tussen de input en output. De relaties tussen input variabelen en de output variabelen worden bepaald door de gegevens, en niet voorbepaald door een model. De volgende gegevens zijn in grafiek 15 gebruikt: tarief, omzet, aantal producten, leeftijd en abonnement. Wanneer deze waarden in het model worden gezet is de juistheid van het model 89% wat betrouwbaar is.



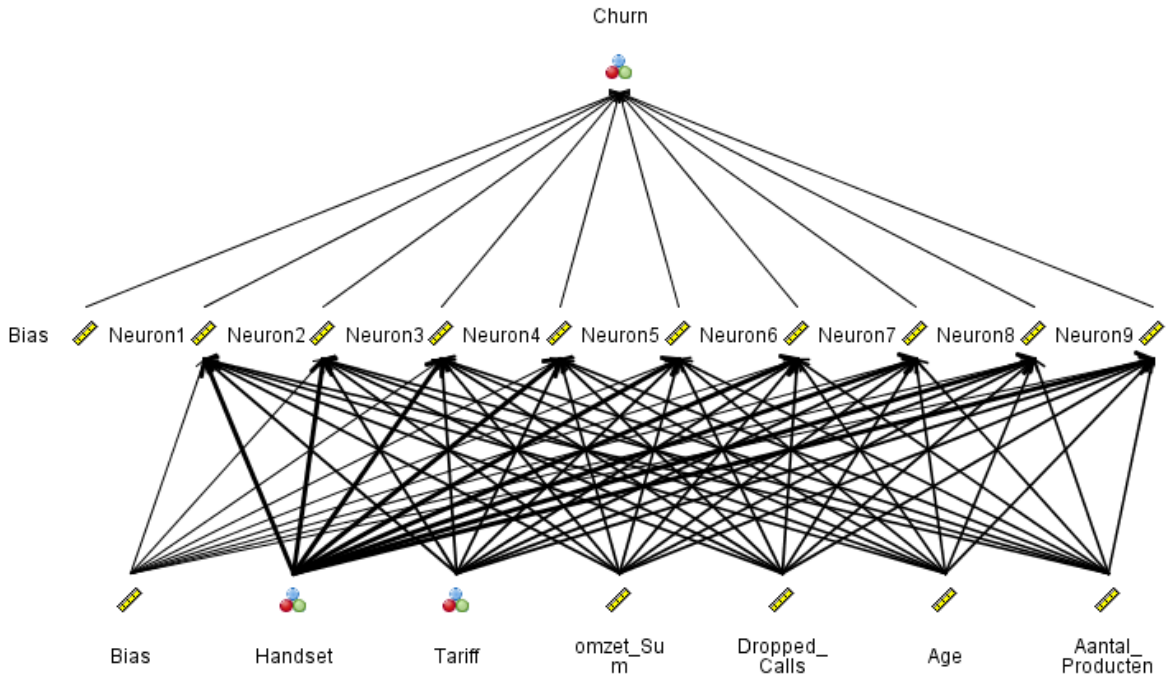
Afbeelding 31 - Juistheid voorspelling churn

Het belangrijkste input veld dat voor deze analyse is geselecteerd is het abonnement. Door middel van dit veld kan de churn het beste voorspeld worden. Vervolgens zijn de overige input velden ongeveer gelijk qua belangrijkheid voor de voorspelling.



Grafiek 15 – Belangrijkheid input velden

In afbeelding 30 wordt het neuraal netwerk weergegeven dat naar aanleiding van de input velden is ontstaan.



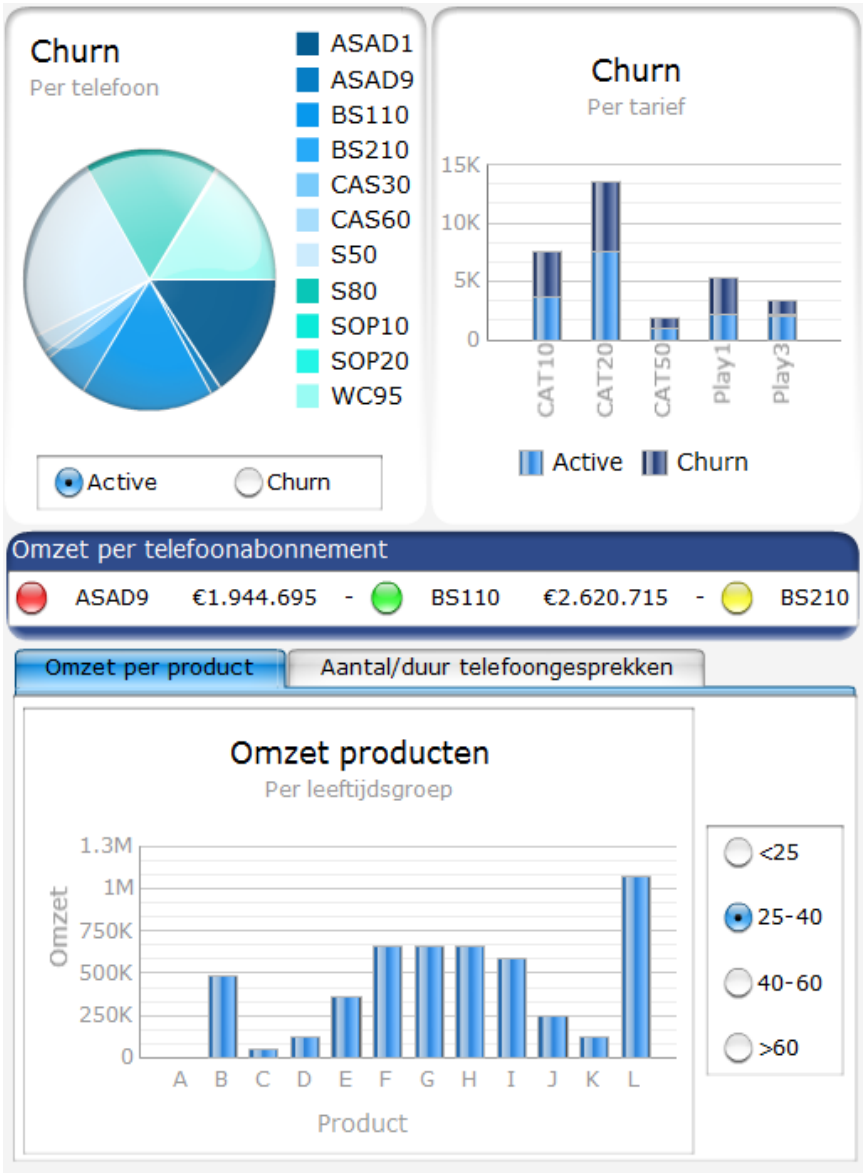
Afbeelding 30 - Neuraal netwerk churn

Bijlage K. Xcelsius

Visualisatie resultaat

Doordat de voorspellingen in de demo voornamelijk gingen over de kans op wel of geen churn was het niet zo interessant om dit in Xcelsius te zetten. Om toch de mogelijkheden van Xcelsius weer te geven zijn er beschrijvende analyses gemaakt van de gegevens van de demo. Door in het begin te inventariseren welke rapportages gemaakt moesten worden, konden de gegevens op de juiste manier worden gerangschikt in Xcelsius. Dit maakte het opstellen van de rapportages wat gemakkelijker en zorgde voor een mooi resultaat.

In afbeelding 32 is de data van de demo in Xcelsius gezet en hiervan zijn onder andere grafieken gemaakt. Hierbij is de churn weergeven per telefoon van een klant en per tarief. Ook is de omzet per telefoonabonnement zichtbaar in een lichtbalk waarbij aangegeven wordt of de target gehaald is van de omzet ('rood=slecht', 'geel=redelijk' en 'groen=goed'). Door middel van Xcelsius kunnen de resultaten in één oogopslag de gegevens inzichtelijk maken waardoor er vervolgens beslissingen kunnen worden genomen. In dit voorbeeld kan er bijvoorbeeld worden geanalyseerd waarom het telefoonabonnement 'ASAS9' qua omzet de doelstelling niet heeft gehaald. Hierop kan worden ingespeeld waardoor het in de toekomst kan verbeteren. In de onderstaande grafiek kan er worden gekeken hoe groot de omzet per product is bij een bepaalde leeftijdscategorie. Aan de hand van dit voorbeeld worden de populaire producten inzichtelijk waardoor deze producten bijvoorbeeld meer kunnen worden gepromoot onder de betreffende doelgroep.



Afbeelding 32 - Resultaat churn verwerkt in Xcelsius

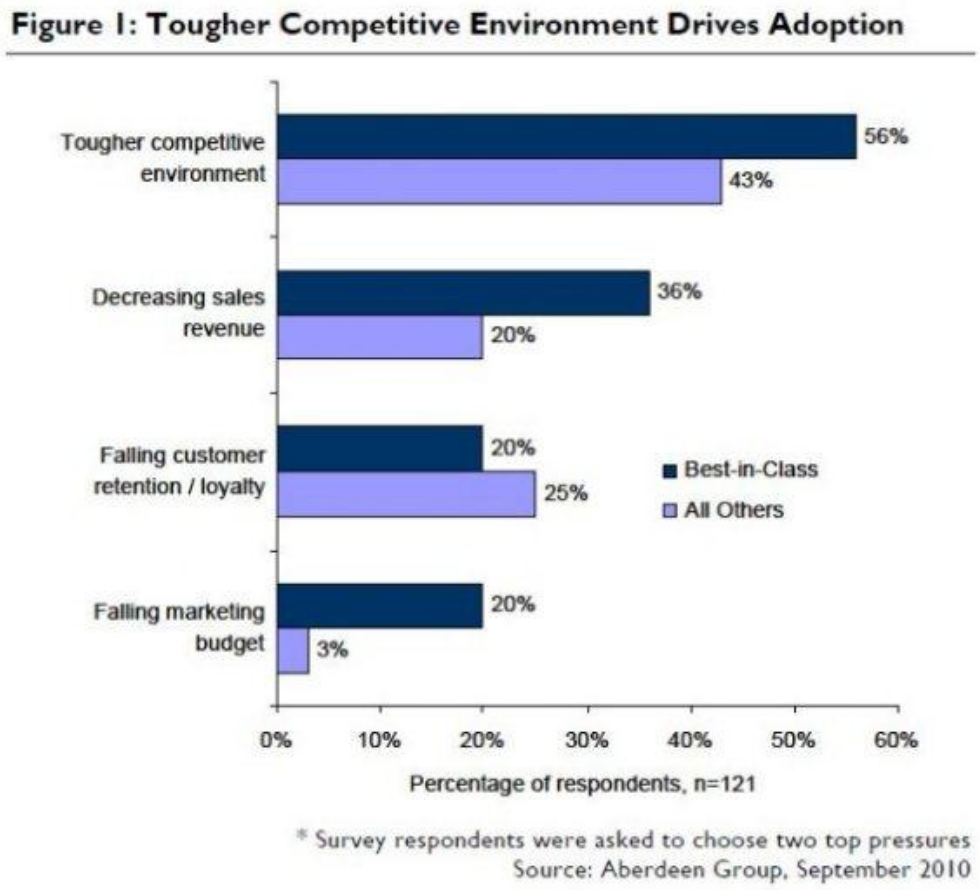
Bijlage L. Artikel

PREDICTIVE ANALYTICS SOFTWARE VAKER GEKOCHT

22/11/2010 Marketeers die werkzaam zijn bij grote, marketinggestuurde bedrijven, schaffen steeds vaker predictive analytics-software aan. Zo vond de Aberdeen Group. De uitkomsten staan in het rapport 'Predictive Analytics – Driving Sales with Customer Insights'.

Het bureau ondervroeg 121 bedrijven. 65 procent van de marketeers die werkzaam zijn bij grote, marketinggestuurde bedrijven, heeft aangegeven tot aankoop over te willen gaan. Bij bedrijven waar marketing een minder belangrijke rol speelt, is dit percentage 36 procent. Volgens de onderzoekers is de steeds zwaarder wordende concurrentie de belangrijkste reden voor aanschaf van predictive analytics-software.

De marketeers van marketinggestuurde bedrijven noemden als tweede belangrijke reden de terugloop van sales (36 procent). Bij bedrijven die niet zo heel erg marketinggestuurd zijn, werden eerst nog tegenvallende retentie/loyalty opbrengsten (25 procent) genoemd, daarna pas tegenvallende sales (20 procent).



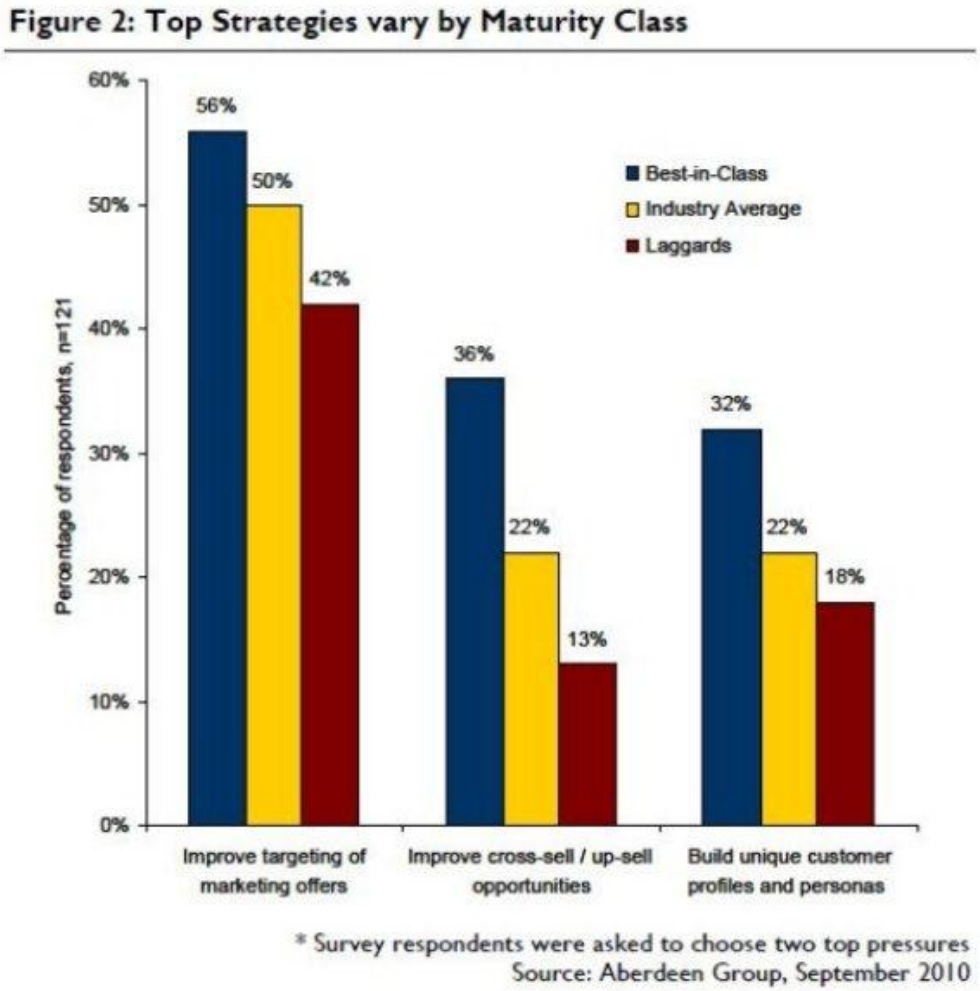
Een ander verschil tussen grote, marketinggestuurde bedrijven en andere bedrijven: 20 procent van de eerste groep overweegt de software te kopen omdat men minder marketingbudget te besteden heeft. In de tweede groep ligt dat percentage op 3 procent.

Best-in-class, gemiddeld en laggard

De onderzoekers verdeelden de bedrijven van de respondenten in drie groepen: best-in-class (20 procent), gemiddeld (50 procent) en laggards (30 procent). Gevraagd naar de belangrijkste uitdagingen antwoordden de marketeers uit groep 1 ten eerste een verbetering van de targeting van marketingaanbiedingen (56 procent), verbeteren van cross- en upselling (36 procent) en het verzamelen en bouwen van unieke klantprofielen en personas.

De andere twee groepen noemden dezelfde uitdagingen, maar in een iets andere volgorde en met iets andere percentages: De gemiddelde groep noemde ook targeting als eerste (50 procent), maar gaf gelijke percentages aan cross- en upselling en de klantprofielen (22 procent). De laggards kozen ook voor

targeting als belangrijkste uitdaging (42 procent), maar kozen vervolgens voor de profielen (18 procent) en daarna de cross- en upselling (13 procent).



Best-in-class respondenten zullen ook eerder een speciaal team samenstellen om nieuwe insights te ontdekken (64 procent vergeleken met 31 procent van de gemiddelde- en de laggards-groep). Volgens de onderzoekers komt dat doordat best-in-class marketeers de beschikking hebben over veel meer klantdata dan hun gemiddelde- en laggard-collega's.

Volgens een ander rapport van Aberdeen Group is 60 procent van de marketeers van grote, marketinggestuurde ondernemingen van plan om de budgetten voor customer analytics en segmenteer-software flink te verhogen in 2011.

Bron: Marketingcharts