

## STUDENT CASE

Stance Lammers - Bachelor Econometrie & Operationele Reserch,  
Master Business Analytics and Quantitative Marketing bij  
Erasmus University Rotterdam

 Data Science  Model Implementatie  Reizen  Volleybal

**underlined**  
Building Data Driven Customer Experience

**Erasmus  
University  
Rotterdam**  
*Erasmus*

### Onderzoeksvraag

Welke modeltechniek kan het beste gebruikt worden om de belangrijkste drivers te vinden van de Net Promoter Score.

### Het doel van Underlined

Underlined biedt middels een drivermodel inzicht aan klanten die willen begrijpen welke factoren invloed uitoefenen op een klant- of bedrijfsmetric (zoals de NPS of churn). **Innovatie** binnen dit product van Underlined zit hem in het **toepassen van technieken die steeds beter scoren op voorspelkwaliteit**. Daarnaast blijft het belangrijk dat de **toegepaste technieken interpreteerbaar** blijven voor onze klanten.

Bijdrage Stance,  
Erasmus Universiteit Rotterdam

### De weg naar het resultaat

Stance is aan de slag gegaan met de opdracht om technieken te vergelijken waarmee Underlined verder kan innoveren op het drivermodel. Zij ging als volgt te werk:

- 1. Literatuuronderzoek:** analyseren van alternatieve modeltechnieken, namelijk ordinale logistische regressie (partial proportional odds model), structure equation modeling en Bayesian networks.
- 2. Preparatie van de data:** het gebruik van technieken om om te gaan met missende waarden in de data.
- 3. Keuze maken:** Focus op twee van de alternatieve modeltechnieken (partial proportional odds model en Bayesian networks).
- 4. Implementatie:** technieken implementeren in R.
- 5. Implementatie extra technieken:** zoals variable selectie en berekenen van 'relative importances'.
- 6. Evalueren en vergelijken:** hoe scoren de technieken waarop gefocust is in de thesis versus de techniek die op het moment door Underlined gebruikt wordt?

### Conclusie

- Zowel het partial proportional odds model als het Bayesian network presteren nóg beter dan het lineaire regressie model op basis van voorspelkwaliteit en interpreteerbaarheid.
- Het Bayesian network presteert het best op basis van voorspelkwaliteit, maar de verschillen in voorspelkwaliteit zijn klein.
- Op basis van interpreteerbaarheid hebben beide modeltechnieken unieke voordelen.
- Door deze unieke voordelen kan niet direct beslist worden welk van deze technieken het beste gebruikt kan worden om de belangrijkste drivers van de Net Promoter Score te bepalen. Er is een vervolg nodig.

## Resultaat

Door middel van de resultaten richt Underlined zich op meerdere mogelijke technieken achter het drivermodel. Op deze manier:

- ✓ Kan per klant worden bekeken welke techniek het beste aansluit op de data.
- ✓ Dit zorgt voor mogelijkheid om nog beter aan te sluiten op de wensen van onze klant; ligt de focus op interpreteerbaarheid of voorspelbaarheid?
- ✓ Kunnen over het algemeen accuratere voorspellingen worden gedaan.

De toepassing van het Underlined drivermodel wordt veelzijdiger waardoor deze nog beter gaat aansluiten bij de situatie van de klant. Ook worden de voorspellingen vanuit het drivermodel nog accurater.

### Informatie

Meer weten over het Studenten Programma van Underlined? Neem dan contact op met Jet of Ana. Kijk op [underlined.nl](https://www.underlined.nl) voor meer informatie.

**underlined**  
Building Data Driven Customer Experience