**Wat is de vraag, of het probleem, waar het paper aan werkt (probeert te beantwoorden of op te lossen)?**

Het paper richt zich op de vraag welke data het meesten kan bijdrage bij het genereren van feedback. Het paper richt zich op het bepalen welke data zoals Learning Management System (LMS) gegevens, e-tutorials, formatieve assessments en leerdisposities de beste voorspellers zijn van prestaties en welke de meest waardevolle feedback kunnen bieden voor studenten en docenten.

**Onderzoeksvragen:**  
**Q1.** In welke mate voorspellen (zelf gerapporteerde) leerhoudingen van studenten, LMS-gegevens en e-tutorialgegevens (formatieve beoordelingen) de academische prestaties in de tijd?  
**Q2.** In welke mate hebben voorspellingen op basis van deze verschillende gegevensbronnen betrekking op unieke aspecten van prestaties, en in welke mate overlappen deze voorspellingen?  
**Q3.** Welke gegevensbron(nen) (leerhoudingen, LMS-gegevens, e-tutorials, formatieve toetsen) bieden de meeste potentie om tijdige feedback aan studenten te verstrekken?

* **Is dat vooral een theoretisch of eerder een praktisch gedreven vraagstuk? Beargumenteer.**

Het is een praktische gedreven onderzoek. Het onderzoek maakt gebruik van data 922 studenten die een wiskunde en statistiek module volgen. Er wordt hierbij data gebruikt uit het Learning Management System, 2 e-tutorials en gegevens van de studenten zelf.

**Welke onderzoeksmethode wordt toegepast? Hoe wordt informatie, data of resultaten verkregen?**

In het onderzoek maken ze gebruik van 102 variabelen afkomstig uit verschillende bronnen:

* **Demografische data:** In het onderzoek gebruiken ze demografische data zoals nationaliteit, geslacht, leeftijd en vooropleiding. Hierbij wordt de laatste omgezet naar naar indicator voor het onderwijs in het Nederlands onderwijssysteem omdat veel van de studenten internationaal zijn. Leeftijd is uiteindelijk uitgesloten omdat het weinig variatie vertoonde. De belangrijkste variabel was het type wiskundetraject in de middelbare school
* **Diagnostische inlaten toetsen:** In week 0 maakte alle studenten en intake toets om inzicht te krijgen op de voorkennis van de student.
* **Leerhoudingen:** De leerhouding wordt onderverdeeld in drie categorieën:
  + **Leerstijl:** Leerstijl wordt gemeten in week 0 en is gebaseerd op vermunt’s model.
    - **Diepgaand leren** (gericht op begrip en samenhang)
    - **Stapsgewijs leren** (systematisch en procedureel)
    - **Concreet leren** (praktijkgericht)
    - **Zelfregulatie** (de student stuurt zelf het leerproces aan)
    - **Externe regulatie** (het leerproces wordt gestuurd door docenten of structuren)
    - **Gebrek aan regulatie** (weinig controle over het leerproces)
  + **Motivatie:** De motivatie wordt ook gemeten in week 0 en is gebaseerd op Motivation and Engagement Wheel.
    - **Adaptieve gedragingen en gedachten** (*Boosters*)
    - **Maladaptieve gedragingen** (*Guzzlers*)
    - **Belemmerende gedachten** (*Mufflers*)
  + **Emoties:** De emoties worden in week 4 gemeten. Zo hebben de studenten genoeg gezien van de module, maar is er ook nog voldoende tijd om de feedback te verwerken. Het is gebaseerd op Achievement Emotions Questionnaire (AEQ)
    - **Plezier in leren** (Enjoyment)
    - **Angst** (Anxiety)
    - **Verveling** (Boredom)
    - **Hopeloosheid** (Hopelessness)
* **Learning management system:** Er werd data verkregen via het Learning management system black board. Hierbij werden trackingsgegevens bijgehouden over zes verschillende leeractiviteiten.
  + Intaketoets
  + Feedback bekijken
  + Zelfrapportages
  + Twee opgenomen colleges per week
  + Praktijkopdrachten en projecten
  + Communicatie
* **E-tutorials:** Er werd data gebruikt uit twee e-tutorials MyMathLab- en MyStatLab. Hierbij werden drie scores geregistreerd.
  + Beheersingsscore
  + Tijd besteed aan de taak
  + Aantal pogingen dat nodig was om het behaalde beheersingsniveau te bereiken.

De werden berekend uit 25 wekelijkse wiskunde vragen en 20 wekelijkse statistiek vragen. Dus zes voorspellers per week.

* **Prestaties:** De prestaties van de student worden meegenomen op basis van quizes en huiswerk.

In het onderzoek gebruiken ze vervolgens een lineaire, hiërarchische regressie model om de data te verwerken.

**Hoe worden de verkregen gegevens geanalyseerd?**

In het onderzoek wordt de date geanalyseerd met behulp van een lineaire, hiërarchische regressie model. Er wordt erbij gekeken naar de invloed van verschillende individuele variabele op de voorspelling. Daarnaast wordt er ook gekeken naar een combinatie van variabel. Hiervoor wordt de correlatie coëfficiënt R gebruikt.

**Welke conclusies worden er getrokken? Zijn die terecht?**

In het paper werd de conclusie getrokken dat toetsing de beste voorspeller is van de prestaties van een leerling. Echter geven ze zelf al aan dat deze niet altijd beschikbaar zijn in het blok en dat wanneer deze er wel zijn dat de resultaten soms te laat binnen zijn. Daarna zijn de gegevens uit de uit de e-tutorial zoals beheersingsniveau, tijdsbesteding en het aantal pogingen de beste voorspeller. Ook de informatie uit de intaketoets en vooropleiding kunnen een goede voorspeller zijn. De tracking op black board had geen toegevoegde waarde.

**Welke discussiepunten bespreekt het paper en/of zie jij zelf?**

* Tracking heeft geen toegevoegde waarde wat eerder onderzoek bevestigd
* Toetsing zijn meestal vertraagde gegevens. Tot wanner deze binnen zijn de gegevens uit de e-tutorial en intaketoets het belangrijkst.
* Het onderzoek richt zich alleen op de formele leer interactie. Het onderzoek houdt geen rekening met andere hulpmiddelen of met studenten die elkaar helpen.
* **Niet in het paper:** Het onderzoek richt zich alleen op de correlaties tussen de data en de prestatie van de studenten, maar kijkt niet naar de causale verbanden. Een correlatie betekend niet dat er ook een causaal verband is. (Zie voor voorbeelden: <https://www.tylervigen.com/spurious-correlations> ) Er zou meer onderzoek gedaan moeten worden of deze voorspellers ook daadwerkelijk invloed hebben op de prestaties.