

# Projectplan stage

Research Neurale Netwerken en Software Cost Estimation

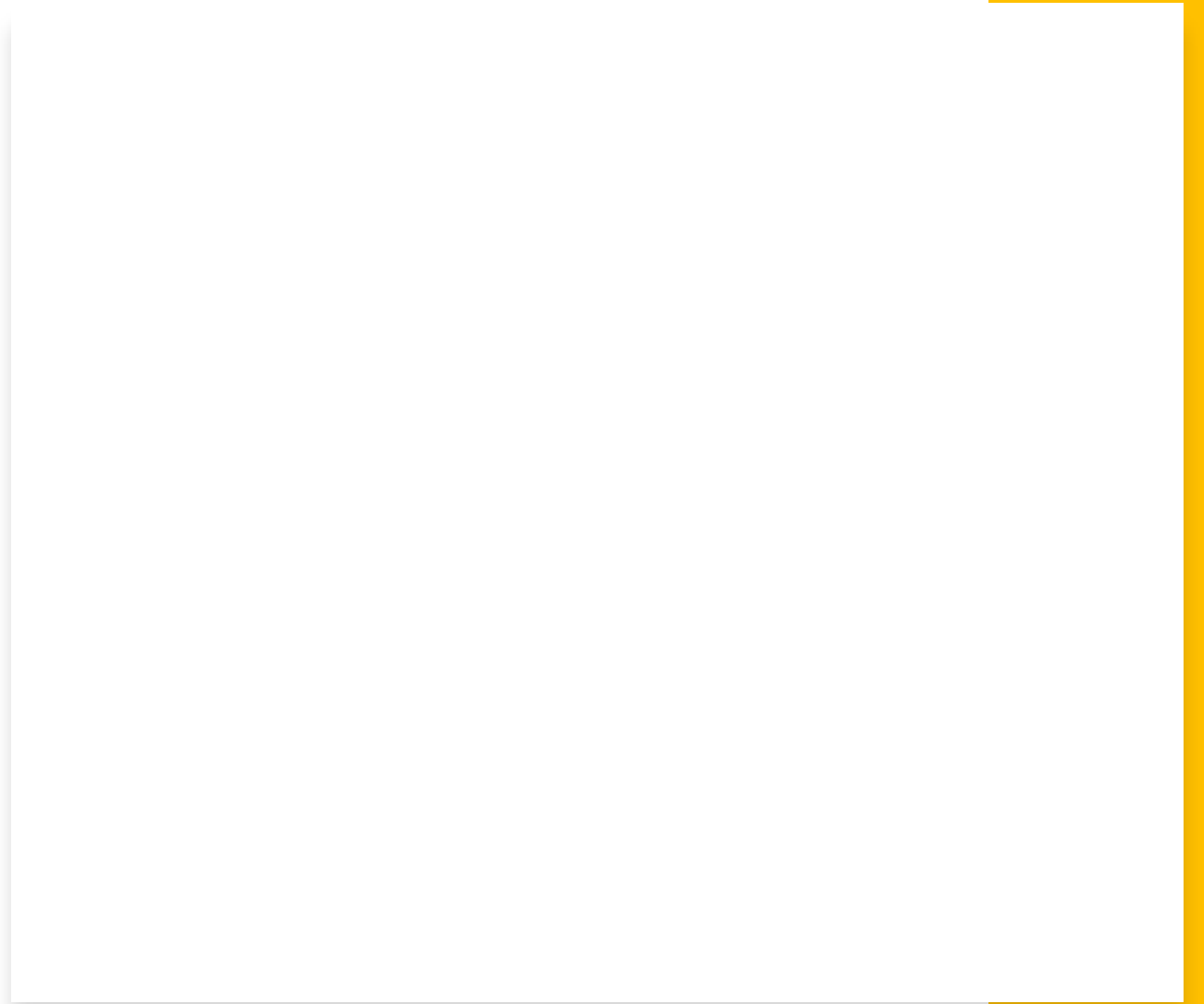
---

# Inhoud

- Achtergrondinformatie & aanleiding
- Verwachte resultaat
- Business case
- Projectafbakening & risicoanalyse
- Informatie & rapportering



# Aanleiding & achtergrond



## Achtergrondinformatie

- Mijn opdracht is voor de Univerzita Tomáše Bati dit is een universiteit in Zlín, Tsjechië.
- Mijn supervisor is Radek Shilhavy, hij is verantwoordelijk voor het verloop van mijn traject.

# Aanleiding project

## Huidige methodes:

- Analogous estimating is een methode die in het projectbeheer wordt gebruikt om de kosten of de duur van een lopend project te schatten door het te vergelijken met een soortgelijk afgerond project.
- Bottom-Up Estimating is een methode die bij projectbeheer wordt gebruikt om de kosten of de duur van een project te schatten door het op te splitsen in individuele componenten en de kosten of duur van elke component te schatten.
- Three-Point Estimating is een methode waarbij voor elk projectonderdeel drie scenario's worden geschat: een optimistische schatting, een pessimistische schatting en een meest waarschijnlijke schatting.

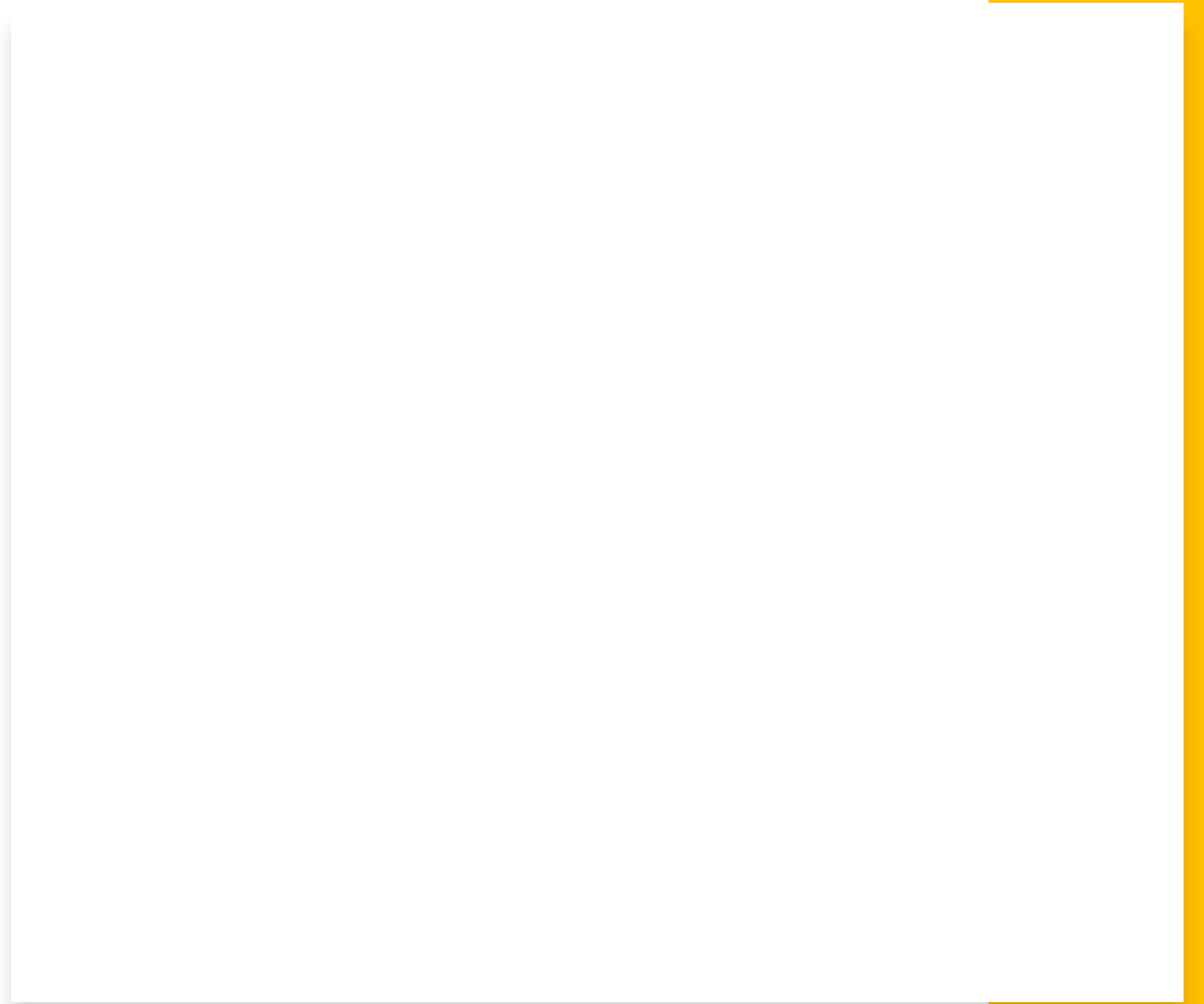
# Aanleiding project

Beperkingen huidige methodes:

- De huidige methodes nemen zeer veel tijd in beslag, er is ook een gebrek aan historische data.
  - Ook zijn er veel menselijke factoren die het process beperken zoals gebrek aan expertise, te veel zelfvertrouwen, vooroordelen etc.
- Ik ga op zoek naar een beter alternatief.  
→ Software Cost Estimation met AI.



Verwachte  
resultaat





# Verwachte resultaat

## **Algemene eisen:**

- AI-modellen die instaan voor de predicties van de geschatte effort van een project, het AI model zal getraind worden op meerdere datasets
  - De datasets zijn gebaseerd op Functional of use-case points.
  - Een Research paper gebaseerd op de experimenten die uitgevoerd zijn.
-





# Verwachte resultaat

Overzicht onderdelen:

- Review paper
  - Datasets cleanen
  - AI ontwerp
  - Research paper
-



# Verwachte resultaat

Onderdelen:

## **Review paper**

- gebaseerd op huidige methodologieën e.g. GEHO-NFN, OLCE, etc.
  - De onderzoeken zijn niet ouder dan 5 jaar → recente technologie
  - Het bestaat uit 3000-6000 woorden
  - Zoeken naar een Gap in knowledge → Deep Learning, weinig data
  - Uitkomst dicteert het verdere onderzoek → research paper
-



# Verwachte resultaat

Onderdelen:

## **Datasets cleanen:**

- Deze datasets bestaan uit uitgebreide historische data over projecten de datasets zijn gebaseerd op functional points of use-case points
  - Er moet onderzocht worden wat de relevante features zijn en hoe deze in relatie met elkaar staan
  - De data moet gecleand worden omdat er zich extreme outliers bevinden binnen de data
  - Hoe beter de data gecleaned is → hoe beter de resultaten
-



# Verwachte resultaat

Onderdelen:

## **AI ontwerp**

- Onderzoek naar geschikt algoritme voor de beste accuraatheid
  - Getraind op meerdere datasets → realistische representatie van het echte werkveld
  - Het resultaat is consistentere dan huidige methodes
  - Kunnen inschatten hoeveel tijd er in een nieuw project zal worden gestoken, zodat de werkkraft optimaal verdeeld kan worden
-



# Verwachte resultaat

Onderdelen:

## **Research paper:**

- Gebaseerd op eigen onderzoek uit de conclusie van de review paper
  - Het bestaat uit 4000-6000 woorden
  - Hier wordt er uitleg gegeven over de experimenten, datasets, algoritmes etc.
-



# Business case & doelgroepen



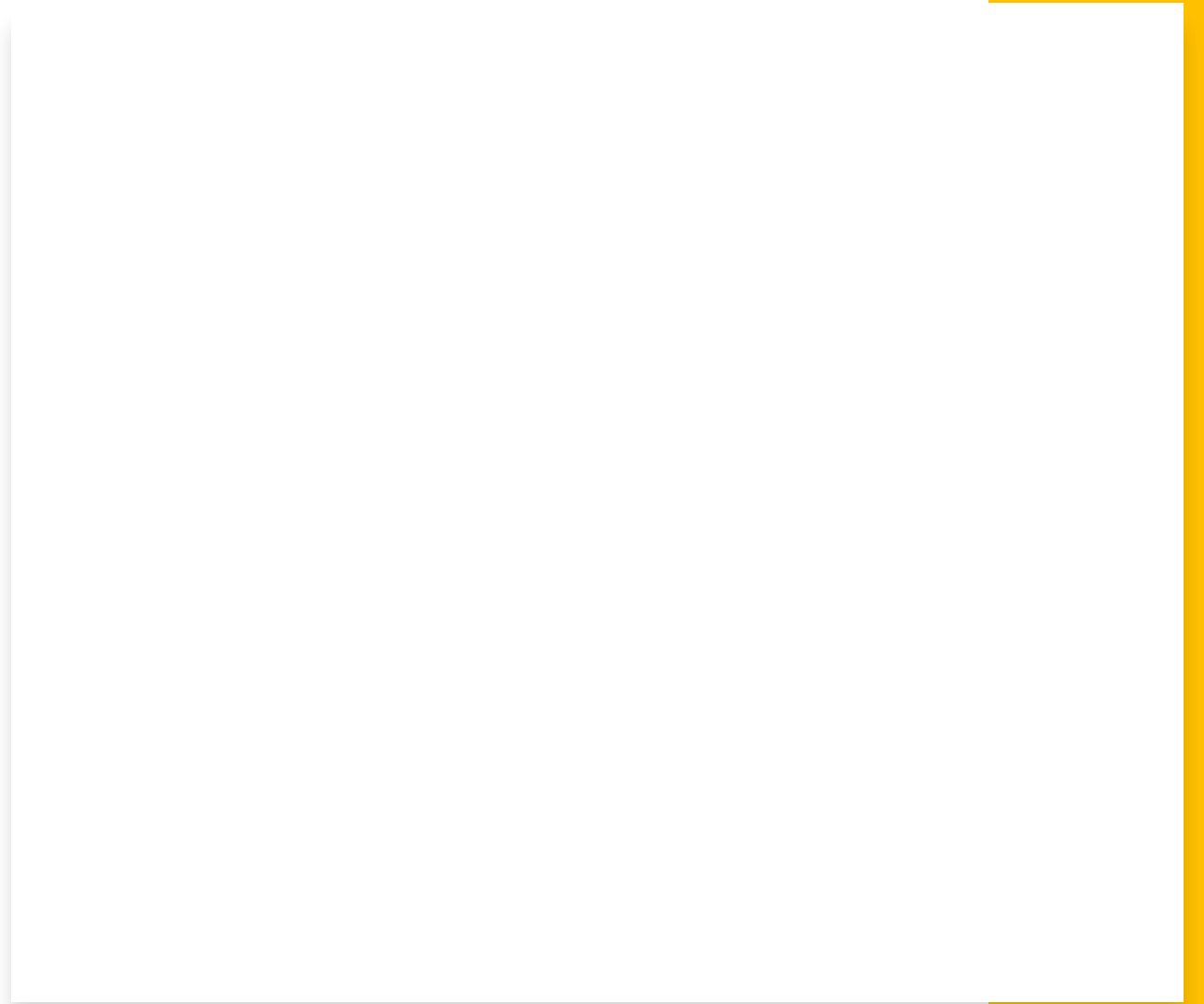
# Business case

## Doelgroepen:

- Fakulta aplikované informatiky utb
    - Doormiddel van het onderzoek te publiceren krijgt de fakultiteit bekendheid, ze hebben ook de mogelijkheid om verder te onderzoeken
  - Software engineering bedrijven | potentiële klant
    - Automatiseren van Software Cost Estimation bespaard veel tijd en kan ervoor zorgen dat een project vlotter verloopt
-



# Projectafbakening & risicoanalyse







# Projectafbakening & risicoanalyse

- Projectafbakening
    - Review paper
    - Experimenteren
    - Research paper
-



# Projectafbakening & risicoanalyse

- Verantwoordelijkheden
    - Opdrachtgever (Prof. Silhavy)
      - Datasets aanleveren zoals de ISBG en China datasets, dit zijn de meest uitgebreide en gebruikte datasets
      - Duidelijke richtlijnen geven
      - Helpen met publiceren van papers naar diverse journals
-




# Projectafbakening & risicoanalyse

- Risico's en maatregelen
    - Weinig code beschikbaar
      - Eigen ervaring en ChatGPT
    - Moeilijke algoritmes
      - Veel onderzoeken
-



# Informatie & rapportering





# Informatie & rapportering

Presentatie projectplan:

- Vrijdag 31/03
  - 11:45
-



# Informatie & rapportering

Inleveren Review paper voor CSOC conference:

- Woensdag 05/04





# Informatie & rapportering

Eerste versie van Research paper inleveren voor review:

- Donderdag 04/05





# Informatie & rapportering

Inleveren van alle documenten voor initiële feedback van supervisor (Thomas More):

- Realisatie document
  - Projectplan
  - Review paper
  - Research paper
    - Vrijdag 19/05
-



# Conclusie

- Duidelijk streefdoel
  - Een gat in de markt vinden
- Veel kennis opdoen
  - Machine learning
  - Deep learning
  - Formeel schrijven

