

CHURN ANALYSE



Forfattere:

Line A. Adolph, Maria B. A. Hitz, Maria Cristiana Maxim, Martin Bindner, Abdikadir A. M. H. Omar

1. interne eksamensprojekt

Vejleder: Simon Bjerrum Eilersen

Dato: 9. maj 2025

Antal tegn: xx.xxx

Indholdsfortegnelse

1	Resumé	3
2	Indledning	3
3	Problemformulering	3
3.1	Underspørgsmål	3
4	Afgrænsning	4
4.1	AI Chatbots	4
4.2	Data	4
4.3	Modeller	5
4.4	Webapplikation	5
4.5	Etik og jura	5
5	Definitioner	6
6	Analyse	6
7	Anbefaling	6
8	Konklusion	6
10	Literaturliste	7
11	Bilagsoversigt	8

1 Resumé

2 Indledning

Business Viborg er registreret under branchekoden 26104793, og arbejder målrettet for at skabe optimale rammer for erhvervslivet i Viborg Kommune. Som en medlemsorganisation med over 600 virksomheder i ryggen, er relationerne til medlemskredsen helt afgørende, både for at dele viden, styrke netværk og skabe lokal vækst. I forbindelse med præsentationen af Business Viborg udtalte chefkonsulent Michael Freundlich: “Vores mål for 2025 er at nå 700 medlemmer og en omsætning på 2,9 mio. kr.”

Men når virksomheder melder sig ud, mister Business Viborg ikke kun en indtægt, men også værdifulde forbindelser, politisk legitimitet og mulighed for at gøre en forskel for erhvervslivet i området. For at handle proaktivt ønsker Business Viborg at få bedre indsigt i, hvad der driver churn og hvem der er i risikozonen.

Derfor skal der udvikles et datadrevet værktøj, som kombinerer teknisk analyse med brugervenlig indsigt. Et værktøj, der gør det muligt for både medlemskonsulenter og ledelse at træffe kloge beslutninger og handle i tide med respekt for både dataetik og jura.

3 Problemformulering

Hvordan kan Business Viborg analysere og anvende medlemsdata til at udvikle et beslutningsunderstøttende dashboard, der forudsiger churn og forklarer centrale risikofaktorer baseret på relevante maskinlæringsmetoder og med inddragelse af etiske og juridiske overvejelser?

3.1 Underspørgsmål

Eksplorativ analyse (EDA)

Beskriv hvilke mønstre og karakteristika kendetegner de virksomheder, der forlader Business Viborg?

Modelvalg og performance

Hvordan kan forskellige machine learning-modeller anvendes til at forudsige churn i Business Viborgs kontekst, og hvilke modeller er mest velegnede?

Datavisualisering

Hvordan kan resultater og churn-indsigter formidles via et brugervenligt dashboard, som understøtter daglig opsøgende indsats for medlemskonsulenter og ledelse?

Etik og jura

Hvilke juridiske krav (fx GDPR) og etiske overvejelser bør indgå i udviklingen og brugen af et churn-forudsigelsesværktøj baseret på medlemsdata?

4 Afgrænsning

4.1 AI Chatbots

I forbindelse med udarbejdelsen af projektet er ChatGPT 4.0 blevet anvendt som et værktøj til grammatisk og sproglig korrektur, idéudvikling samt forbedring af det sproglige udtryk. Modellen har udelukkende fungeret som et hjælpemiddel i arbejdet med tekstbaserede opgaver og har hverken erstattet selvstændig analyse, kritisk refleksion eller været brugt til direkte besvarelse af problemformuleringen

I dette projekt fokuserer vi på at udvikle en webbaseret applikation, der kan forudsige medlems-churn og visualisere centrale churn-drivere for Business Viborg. Projektet indebærer derfor flere afgrænsninger.

4.2 Data

Vi anvender udelukkende datasættet, som er udleveret af Business Viborg. Datasættet indeholder information om medlemskab, virksomhedsstørrelse, branche, engagementsniveau, kontaktfrekvens og deltagelse

i arrangementer. Dataene er anonymiserede og begrænset til et specifikt tidsrum, hvilket kan påvirke modellens generaliserbarhed.

4.3 Modeller

Projektet fokuserer på udvikling og afprøvning af machine learning-modeller til churn-prediktion, herunder blandt andet Random Forest, XGBoost og Lasso. Vi anvender bootstrapping som resampling-metode og benytter AUC og accuracy som de primære evalueringsmål. Fokus er på modellens forklaringskraft og praktiske anvendelighed frem for at optimere for den højest mulige præcision.

4.4 Webapplikation

Applikationen er en prototype og er ikke integreret med Business Viborgs IT-systemer eller databaser. Den kan uploades og anvendes lokalt eller via RStudio Cloud.

4.5 Etik og jura

Vi behandler kort de etiske og juridiske aspekter ved brugen af persondata i overensstemmelse med GDPR, men foretager ikke en fuld juridisk gennemgang. Implementering af samtykke- og adgangsstyring indgår ikke som en del af denne prototype.

5 Definitioner

6 Analyse

7 Anbefaling

8 Konklusion

9

10 Literaturliste

AI

OpenAI. (2025). ChatGPT (4.0). <https://chatgpt.com/>

Bøger

WWW-dokumenter

Undervisningsmaterialer

11 Bilagsoversigt

- Bilag 1: x
- Bilag 2: x
- Bilag 3: x
- Bilag 4: x
- Bilag 5: x
- Bilag 6: x

1. Data load

```
# Data load
```

2. Merge datasets

```
# Merge datasets
```

3. Clean data

```
# Clean data
```

4. Feature Engineering

Feature Engineering

5. EDA

EDA

6. Preprocessing

```
# Preprocessing test test testtest test test test test test test test test test test
```

7. Modelling

```
# Modelling
```

8. Evaluate metrics

```
# Evaluate metrics
```