**Business Intelligence og dataanalyse til data-drevne beslutninger i GAME ngo**

Hovedopgave – Pelle Hald Vedsmand

Contents

**No table of contents entries found.**

1. Forord

Efter mit praktikophold hos GAME i perioden 5/8 – 28-11, har jeg været i samarbejde om at udvikle et produkt for firmaet.  
**Udvikling i praktikophold**:  
Under min praktikperiode skabte jeg en pipeline, der skulle automatisere og forbedre GAMEs arbejde med dataanalyser. Derudover hjalp jeg forskellige teams i GAME med analyse og modellering af data, dog på projektbasis.

1. Indledning

Hvordan kan data-drevne beslutninger skabe success for GAME? Og hvordan kan data engineering samt Business Intelligence bruges til at skabe disse beslutninger?  
Dette var den overordnede problemstilling i virksomheden.   
Denne rapport vil beskrive, hvordan man kan bruge Business Intelligences principper og værktøjer til at skabe analyser og modeller, som vil kunne påvirke GAMEs vækst som firma.

Under min praktikperiode håndterede jeg dataanalysearbejde i flere projekter. Heri fik GAME en forståelse for mulighederne med Business Intelligence.  
  
I følgende rapport er der opstillet forskellige prediction models, baseret på data i og udenfor GAME, der skal gøre det muligt at forudsige deltagelsen til fremtidige events.  
GAME vil med denne viden kunne optimere events med flere/færre instruktører og materiel. De vil også få dannet en større forståelse for diverse deltagelsesmønstre, samt korrelationen mellem deltagelsestal og features, både i GAME, men også udefrakommende.

1. Virksomhed

GAME er en ngo der forsøger at skabe social forandring hos børn og unge igennem aktiviteter og sport. GAME træner unge frivillige til at være instruktører og rollemodeller[[1]](#footnote-1)

GAME er en NGO, der er sponsoreret af diverse fonde, firmaer og legater. GAME skal derfor ud at bevise for investorerne, at de har formået at løse de opgaver der stilles.

Udviklingen i GAME er meget flydende i GAME, dvs. at de ikke følger nogle stringente modeller eller metoder. Det meste af arbejdet foregår på projektbasis, gennem investorer.  
Disse projekter ledes af flere forskellige teams, der alle har brug for analyser af indkommende data.

**Formål:**

Udviklingen af prediction models vil hjælpe GAME med at vokse som virksomhed, da de vil kunne skabe bedre finansielt råderum, ved at kunne distribuere budgettet til de rigtige events.

**Problemstillinger:**

* GAME har problemer med at forudsige antal deltagere til fremtidige events, og de bruger derfor, til tider, for mange penge på events med lav deltagelse.
* GAME har svært ved at forstå hvilke ”features”, der påvirker deres deltagelsestal til events.

**Mål og resultater:**

1. Krav

**User Stories**

**User Story 1: Forudsigelse af deltagertal**

* **Som en** projektleder hos GAME
* **vil jeg** kunne forudsige deltagertal for fremtidige events for Playmaker Teamet
* **så jeg kan** allokere de rette ressourcer (instruktører og materialer) effektivt.
  + **Acceptance Criteria:**
    1. En model skal kunne estimere deltagertal baseret på historiske data med mindst 85% nøjagtighed.
    2. Output skal præsenteres som en rapport eller et dashboard, der kan tilgås af projektledere.
  + **Estimate:** Large (Flere modeller, baseret på data i GAME og udefrakommende data).

**User Story 2: Identifikation af vigtige faktorer**

* **Som en** projektleder hos GAME
* **vil jeg** identificere de vigtigste faktorer, der påvirker deltagelse i events
* **så jeg kan** give anbefalinger til optimering af fremtidige arrangementer.
  + **Acceptance Criteria:**
    1. En liste over de top 5 vigtigste features (fx vejr, ugedag, eventtype) skal præsenteres med visualiseringer.
    2. Data skal være dokumenteret, så metoderne kan gentages.
  + **Estimate:** Medium (Dataforberedelse og kørsel af modeller som fx feature importance i Random Forest).

**User Story 3: Analyse af historiske data**

* **Som en** analytiker hos GAME
* **vil jeg** kunne analysere historiske deltagelsesmønstre
* **så jeg kan** finde tilbagevendende tendenser, der kan bruges til planlægning.
  + **Acceptance Criteria:**
    1. Analyse af deltagelsesmønstre skal inkludere visualiseringer af sæsonvariationer.
    2. Resultaterne skal præsenteres i et læsbart format for ikke-tekniske brugere.
  + **Estimate:** Small (Eksisterende data og basisanalyse).

1. Valg af teknologi(er)
2. Valg af arkitektur
   * 1. Python
     2. Scikit-learn
3. Design
4. Implementering
5. Test
6. Evaluering / konklusion

1. <https://game.ngo/what-we-do/> [↑](#footnote-ref-1)