# Mallintaminen

Mallintaminen on CRISP-DM:n neljäs vaihe, ja sillä tarkoitetaan muokatun ja puhdistetun datan mallintamista malliksi, joka täyttää tiedonlouhinnan tavoitteet ja vastaa liiketoiminnan asettamiin tavoitteisiin. Mallinnuksella tarkoitetaan esimerkiksi koneoppivan, tai syväoppivan algoritmin opettamista datasta. Opetetulla mallilla voidaan ennustaa haluttua tähtäinmuuttujaa.

## Mallinnustekniikka

Kuvaa alle mahdollisimman tarkasti käytettävä mallinnustekniikka ja mallinnuksen ominaisuudet

|  |
| --- |
|  |

## Mallinnuksen olettamukset

Kuvaan alle kaikki olettamukset, joita valitsemasi mallinnustekniikka tekee. Jotkut mallinnustekniikat vaativat esimerkiksi, ettei datassa ole puuttuvia arvoja, tai sen käyttämä datan jakauma asettuu tietylle välille. Kaikki tällaiset olettamukset tulee kirjata ylös.

|  |
| --- |
|  |

## Testaussuunnitelma

Kuvaa alle mallin koulutuksen-, testauksen- ja validoinnin suunnitelma.

|  |
| --- |
|  |

## Malliraportti

Kuvaa alle luomasi malli, sisältäen läpikäynnin mallin ominaisuuksista, sekä sen rakennusvaiheista. Jos mallinnuksen aikana on vastaan tullut ongelmia kuvaa myös ne ja niiden ratkaisut. Sisällytä kuvaukseen myös mallinnuksessa käytetyt parametrit ja valintaperusteet kyseisille parametreille.

|  |
| --- |
|  |

## Mallin arviointi

Listaa alle kaikki luomasi mallit, niiden ominaisuudet, kuten tarkkuus, koulutusaika ja mahdolliset virheet. Vertaa niitä keskenään, sekä järjestä mallit paremmuus järjestyksessä valittujen arviointi kriteerien mukaan. Jos jatkat mallin parantelua esimerkiksi hyperparametrioptimoinnilla, raportoi myös tehdyt muutokset ja niiden vaikutukset.

|  |  |
| --- | --- |
| Valittu arviointi kriteeri |  |

### Mallit, kierros 1

|  |  |  |
| --- | --- | --- |
| Malli | Suoritusteho | Mahdollinen optimointi |
|  |  |  |
|  |  |  |
|  |  |  |

### Mallit, kierros 2

|  |  |  |
| --- | --- | --- |
| Malli | Suoritusteho | Mahdollinen optimointi |
|  |  |  |
|  |  |  |
|  |  |  |

### Mallit, kierros 3

|  |  |  |
| --- | --- | --- |
| Malli | Suoritusteho | Mahdollinen optimointi |
|  |  |  |
|  |  |  |
|  |  |  |