



alaviiva

Metropolia Ammattikorkeakoulu

Tieto- ja viestintätekniikka

Datan käsittely ja koneoppiminen

16.3.2021

Etunimi Sukunimi – Etunimi Sukunimi – Etunimi Sukunimi

Analyysiraportti (CRISP-DM)

Tehtävä n:o 1

*Korvatkaa alla kohdissa 1-6 olevat ohjetekstit omalla tekstillänne. Voitte käyttää selventäviä kuvia ja ruutukaappauksia. Päivittäkää kansisivulle tekijöiden nimet ja tehtävän numero.*

# Tavoitteet

Mikä on tämän koneoppimis-/tiedonlouhintaprojektin tavoite organisaation tai liiketoiminnan näkökulmasta? Mihin kysymyksiin projektissa pyritään vastaamaan? Mitä yritetään saada selville?

# Data

Miten data on tuotettu? Kuvatkaa jokainen attribuutti: minkä tyyppinen attribuutti on kyseessä (luokkamuuttuja, numeerinen muuttuja jne.) Mitkä ovat arvojen vaihteluvälit? Onko puuttuvia tai poikkeavia arvoja? Mitä mahdollisia ongelmia ja virhelähteitä dataan liittyy? Mikä on datan koko ja mitä siitä voi päätellä?

# Datan valmistelu

Mitä esikäsittelytoimenpiteitä teitte datalle ja miksi? Kuvatkaa toimenpiteet niin yksityiskohtaisesti, että ne ovat toistettavissa.

# Mallinnus

Miten analysoitte datan? Mitä tiedonlouhintamenetelmiä sovelsitte ja millaisin parametrein? Miksi valitsitte juuri kyseisen menetelmän? Millaisia tuloksia saitte? Esittäkää keskeiset tulokset sellaisinaan (kuvat / taulukot). Pelkkä tulosten sanallinen kuvailu ei riitä.

# Arviointi

Miten tulkitsette ja selitätte tulokset? Onko tuloksissa jotain yllättävää ja selittämätöntä?

# Täytäntöönpano

Mitä tulokset tarkoittavat tavoitteiden näkökulmasta? Miten vastaatte ensimmäisen kohdan kysymyksiin? Mikä on toimenpidesuositus, jos sellaisen voi laatia?