

Pikaruokaravintolan asiakasvirran ennustaminen/mallien vertailu (aihe-esittely)

Henrik Purokoski

13.6.2025

Ohjaaja: *Oscar Salonaho*

Valvoja: *Pekka Pere*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Taustaa

- Asiakasvirran ennustamisesta on hyötyä monessa ravintolan kriittisessä toiminnossa (vuorosuunnittelu, varastohallinta)
- Tällä hetkellä ennustusmallien käyttö on vähäistä ja ennusteet tehdään suurimmaksi osaksi arvaamalla
- Analyttiset mallit antavat yleensä parempia tuloksia kuin pelkkä arvaaminen
- Sopiva malli vähentäisi hävikkiä ja johdon työtaakkaa ja parantaisi myynti-arvioita
- Eri mallien vertailu avartaa näkökulmaa

Box-Jenkins –menetelmä (ARMA)

- Matemaattinen menetelmä aikasarjojen mallintamiseen ja ennustamiseen
- Tutkitusti toimiva ja tarpeeksi yksinkertainen
- Kaksi osaa: autoregressiivinen (AR) ja liukuva keskiarvo (MA)
- Malliin voidaan myös lisätä kausivaihtelu sekä differointi (SARIMA)
- Myös ulkoiset muuttujat mahdollisia (SARIMAX)

Koneoppimismallit

- XGBoost
 - Tehokas koneoppimismalli
 - Helppo soveltaa aikasarjojen ennustamiseen
 - Olemassa olevat kirjastot
 - Python
 - R
- Prophet
 - Aikasarjojen ennustamista varten luotu koneoppimismalli
 - Kausittaisuus huomioitu
 - Olemassa olevat kirjastot

Asiakasmäärän arviointi

- Yksi lasku ei välttämättä aina merkitse yhtä asiakasta
- Nimenomaan asiakkaiden määrän arviominen hyödyksi
- Asiakasmäärän arviomiseen on hyvä kokeilla erilaisia parametrerejä
 - Aterioiden määrä
 - Juomien määrä
 - Ranskalaisten määrä
 - Tilauksen kokonaishinta

Tavoitteet

- Muokataan annettua dataa niin, että se pystyy jollain tasolla arviomaan asiakkaiden määrää
- Luodaan ja parametrisoidaan aikasarjalle sopiva SARIMA –malli
- Testataan mallin hyvyyttä erilaisilla diagnostisilla testeillä
- Verrataan mallia ainakin kahteen eri koneoppimismalliin sekä yksinkertaiseen, arvaamista simuloivaan malliin (esim. Edellisen viikon saman päivän asiakasmäärä)

Rajaukset

- Työn pitämiseksi tiiviinä ainoastaan SARIMA –mallia analysoidaan tarkasti, muut mallit toimivat enemmänkin vertailukohtina.
- Datana käytetään vain osaa ravintoloista, eikä mallia luoda jokaiselle ravintolalle erikseen
- Mahdollisen ennustemallin soveltaminen jää jatkotutkimusten aiheeksi

Menetelmät/työkalut

- SQL datan kyselyyn tietokannasta
- Python aikasarjan ennustamiseen

Tietolähteet/aineistot

- Sclove et al.: Data mining on time series: an illustration using fast-food restaurant franchise data (2001)
- Taylor et al.: Forecasting at Scale (2017)
- Chen et al.: XGBoost: A Scalable Tree Boosting System (2016)

Aikataulu

- Aiheeseen tutustuminen 06/2025
- Aiheen esittely 06/2025
- Työn kirjoittaminen 06-09/2025
- Mallien kehittäminen 07-09/2025
- Valmis työ 10/2025