

# Toteutusdokumentti

## Ohjelman yleisrakenne

Ohjelmassa tärkeimmät luokat ovat:

AbstractElement	Käsittää ”yksikön” sovelluksessa.
Item <i>extends</i> AbstractElement	Määrittää tuotteen, jolla voi olla eri kategorioissa, sillä voi olla erilaisia ominaisuuksia ja sillä on arvosteluja, jotka käyttäjät ovat tehneet.
Quality <i>extends</i> AbstractElement	Määrittää ominaisuuden, joka voi olla eri kategorioissa. Se sisältää tuotteita ja sille voidaan antaa painoarvo, joka kertoo kuinka tärkeä ominaisuus se on tuotteen kannalta.
Rating <i>extends</i> AbstractElement	Määrittää arvostelun. Arvostelulla on aina luoja, arvosteltu tuote ja arvosana.
User <i>extends</i> AbstractElement	Määrittää käyttäjän, jolla on arvosteluja.
Star	Määrittää arvosanan, joka on annettu arvostelussa.
Value	Määrittää painoarvon, joka kertoo kuinka tärkeä ominaisuus on tuotteen kannalta.
SimilarUser	Parittaa käyttäjän (luokan User yksikön) ja sen samuusarvon yhteen olioon, jotka on myöhemmin helppo järjestää.
RecommendedItem	Parittaa tuotteen (luokan Item yksikön) ja sen suositteluarvon yhteen olioon, jotka on myöhemmin helppo järjestää.
TestDataGenerator	Algoritmin testaukseen käytettävän datan generoija.
Example	Luokka, jossa havainnollistetaan Algorithm luokan käyttöä
Algorithm	Itse algoritmi.

## Ohjelman aikavaativuudet

Mielenkiintoisten käyttäjien löytäminen käyttäjälle on aikavaativuudeltaan  $O(ru)$ , missä  $r$  on käyttäjän tekemien arvostelujen määrä ja  $u$  on käyttäjien määrä jotka ovat arvostelleet saman tuotteen kuin  $r$ .

Mielenkiintoisten tuotteiden löytäminen on  $O(u)$ , missä  $u$  on käyttäjien määrä.

Käyttäjien samanlaisuuden määrittäminen on  $O(ur)$  missä  $u$  on vertailtavien käyttäjien määrä ja  $r$  on vertailtavien käyttäjien arvostelujen määrä.

Parhaimman tuotteen löytäminen käyttäjälle  $x$  on aikavaativuudeltaan  $O(ur)$ , missä  $u$  on samankaltaisten käyttäjien joukko ja  $r$  on samankaltaisen käyttäjän tekemien arvostelujen määrä.

# Algoritmin toiminta

## Recommend Items

Suosittelee tuotteita muiden käyttäjien arvostelujen perusteella. Parhaiden tuotteiden suosittelu on jaettu kolmeen osaan, jotka ovat seuraavat:

1. Löydä käyttäjät, jotka ovat meidän käyttäjällemme relevantteja. Tässä tapauksessa otamme kaikki käyttäjät, jotka ovat arvostelleet ainakin yhden saman tuotteen kuin meidän käyttäjämme
  - **Algoritmin toiminta:** Jokaisesta tuotteesta luetaan kaikki arvostelut ja lisätään arvostelijat uuteen listaan, joka palautetaan.
2. Määrittele kuinka samankaltaisia meidän käyttäjämme on kohdan yksi (1) käyttäjien kanssa ja ota tuloksesta paras 10%.
  - **Algoritmin toiminta:** Katso jokainen käyttäjä ja vertaa käyttäjän tekemiä arvosteluja meidän käyttäjäämme siten, että jos arvosteltu tuote on sama niin lisää samuusarvoksi  $5 - \text{abs}(a-b)$ , jossa  $a$  on meidän käyttäjämme arvio ja  $b$  on vertailtavan käyttäjän arvio. Esimerkiksi jos  $a=5$  ja  $b=4$  niin samuusarvo on 4. Tämä toistetaan jokaiselle yhteiselle arviolle ja lasketaan tulos yhteen. Käyttäjä on sitä samankaltaisempi mitä suurempi loppuarvo on. Lopuksi käyttäjät järjestetään sen mukaan kenellä on suurin samuusarvo ja näistä käyttäjistä otetaan paras 10%.
3. Suosittele kohdan kaksi (2) tuotteita perustuen kohdan (3) käyttäjien arvosteluihin.
  - **Algoritmin toiminta:** Katsotaan jokaisen samankaltaisen käyttäjän arviot ja lasketaan suositteluarvo jokaiselle arvostellulle tuotteelle. Suositteluarvo on yhteenlaskettu arvo jokaisen vertailtavan käyttäjän henkilökohtaisesta suositteluarvosta kyseiselle tuotteelle. Henkilökohtainen suositteluarvo on kyseisen henkilön antamat tähdet kyseiselle tuotteelle yhteenlaskettuna tämän käyttäjän samuusarvoon. Lopuksi tuotteet järjestetään sen mukaan millä on suurin suositteluarvo ja lopullinen annetaan vastaus.

## Highest Rated Items

Suosittelee tuotteita niiden keskiarvoisen arvosanan mukaan. Esimerkiksi jos tuotteella A on arvosteluja 1, 3, 4, 4, 5 niin tällöin tuotteen A keskiarvosana on  $(1+3+4+4+5)/5 = 3,4$ . Tuotteet järjestetään niin, että korkein keskimääräinen arvosana on ensimmäinen.

## Similar Items

Suosittelee tuotteita sen perusteella kuinka samankaltaisia ne ovat verrattuna annettuun tuotteeseen.

Tuotteiden samuus lasketaan vertaamalla jokaista tuotetta annettuun tuotteeseen, jokainen yhtäläinen ominaisuus (Quality) lisää tuotteen samuusarvoa niiden ominaisuuksien painoarvon verran. Esimerkiksi jos tuotteella A on ominaisuus X ja tuotteella B on ominaisuus X. Lisäksi tuotteelle A ominaisuus X on painoarvoltaan 3 kun taas tuotteelle B ominaisuus X on painoarvoltaan 4. Tällöin niiden ominaisuuksien arvo on  $(3+4)/2 = 3.5$ . Nämä painoarvot lasketaan yhteen niin saadaan tuotteen kokonaisarvo. Lopuksi tuotteet järjestetään suurimmasta samuusarvosta pienimpään ja tulostetaan.

## Työn mahdolliset puutteet ja parannusehdotukset

Puutteet: Järkevä tapa käyttäjälle lisätä omaa dataa kätevämmiin kuin hard-koodaamalla.

Mahdollisuus filtteräidä tuloksia kategorian perusteella.