Изготвили:

Сияна славова, 24963

Иван Капукаранов, 24958

Проект за курса „Препоръчващи системи“

Задача: recSys 2013 challenge

Съдържание

[Проблем 1](#_Toc454038202)

[Съществуващи решения 1](#_Toc454038203)

[Избрано решение 1](#_Toc454038204)

[Експерименти 1](#_Toc454038205)

[Резултати 1](#_Toc454038206)

[Използвана литература 1](#_Toc454038207)

# Проблем

Данните за задачата са предоставени от [Yelp](http://www.yelp.com/), като съдържат над:

* 10,000 фирми (business)
* 8,000 check – in sites
* 40,000 потребителя (users)
* 200,000 ревюта (reviews)

Трябва да се предскаже рейтинга, който потребител ***u*** би дал на продукт ***i***. За целта са дадени обучаващо и тестово множество, като тестовото множество съдържа само колони потребител и фирма.  
  
Оценяването на алгоритъма става чрез използването на средно квадратична грешка.

# Съществуващи решения

# Избрано решение

За реализиране на задачата използваме [Apache Mahout](http://mahout.apache.org/). За предварителна обработка на данните и оценяване на алгоритмите използваме [R](https://www.r-project.org/).

# Експерименти

Results.txt (може да се направят като таблица)

# Резултати

Resources/graphics/parameter\_change – графики, най –добро при preference 8 user based ~1.05

# Използвана литература

* Идея: <https://www.kaggle.com/c/yelp-recsys-2013>
* Apache Mahout: <http://mahout.apache.org/>