

Recommender System für REKA Geld Annahmestellen

Studiengang: MAS | Vertiefung: MAS Data Science

Die REKA Card ist bereits heute ein weit verbreitetes Zahlungsmittel und kann an über 9000 Annahmestellen in der Schweiz verwendet werden. Die meisten Kunden verwenden die REKA Card nur um damit Bahntickets oder Abos zu kaufen oder ihr Auto zu tanken. Die REKA möchte nun ihre Kunden proaktiv mittels eines Recommender Systems über Annahmestellen informieren, um so den Mehrwert einer REKA Card noch besser ausweisen zu können.

Ausgangslage

Die momentane Ausgangslage ist, dass ein REKA Kunde aktiv nach Annahmestellen für REKA Geld suchen muss. Der Kunde erfährt nicht, wenn in seiner Umgebung eine neue Annahmestelle vorhanden ist (siehe REKA Guide App). Um die Multibrandkarte REKA Card attraktiver zu machen, soll der Kunde in Zukunft proaktiv mittels Recommendation auf Annahmestellen aufmerksam gemacht werden. Als zweiter Punkt soll die Verbreitung der Annahmestellen verbessert werden. Durch den Umstand, dass dem REKA Kunden aktiv Annahmestellen vorgeschlagen werden, wird deren Sichtbarkeit deutlich besser. Somit steigt für Annahmestelle die Attraktivität die REKA Card als Zahlungsmittel zuzulassen. Mit dieser Master Thesis soll nun ein Recommender System umgesetzt werden, welches als Proof of Concept dienen soll in Zukunft vermehrt auf Personalisierung und Datenanalyse zu setzen.

Umsetzung

Für die Umsetzung wurden vier Zielsetzungen definiert, welche innerhalb der Thesis erreicht werden sollen:

- Datenanalyse
- Umsetzung Recommender System
- Umsetzung einer Near Real Time Model Training mit Kafka und Spark
- Visualisierung der Recommendations mittels Web Service und einer Mobile Application

Im Fokus der Ziele stand, den vollständigen Machine Learning Prozess umzusetzen. Das heisst, dass von der Datenanalyse bis zum inkrementellen Training und der Visualisierung alles umgesetzt wird. Es wurden

verschiedene Modelle erstellt, die einen mit der «Surprise Library» und die anderen mittels Tensorflow und Keras.

Diese Modelle wurden bewertet. Für die Bewertung wurden die folgenden bekannten Metriken in Recommender Systemen eingesetzt, Mean Absolute Error, Root Mean Squared Error, Hit-Rate, Novelity und Diversity. Anhand dieser Metriken sind zwei Modelle ausgewählt worden die ähnlich guten Ergebnisse erzielten, das eine ist ein Item-based Ansatz mittels KNNBasic und der Kosinus Distanz und dass andere ein neuronales Netzwerk mit einem autoencoder Aufbau.

Fazit

Innerhalb der Umsetzung der Thesis konnte ich das erarbeitete Wissen aus allen CAS Modulen einsetzen. Von der Datenanalyse, Machine Learning bis hin zum Aufbau einer Trainings-Umgebung. Es konnten zwei gute Modelle erstellt und in die Trainings-Umgebung integriert werden. Die grösste Herausforderung war der Aufbau der Trainings-Umgebung, die Surprise Library ist nicht geeignet für ein verteiltes Training innerhalb von Spark. Anstelle des Trainings innerhalb von Spark auszuführen, wurde ein CRON Job erstellt der täglich ausgeführt wird. Schlussendlich konnte der Proof of Concept erfolgreich umgesetzt werden und der Beweis erbracht werden, dass es möglich ist, REKA Card Annahmestellen den REKA Kunden vorzuschlagen.



Oliver Hofmann

```
Top 10 recs
Item: 9000-ST. GALLEN-Coop Mineralöl, Estimated Rating: 5
Item: 9000-ST. GALLEN-Direkter Verkehr (ohne SBB), Estimated Rating: 4.796792440791675
Item: 9014-ST. GALLEN BRUGGEN-SBB Stationen, Estimated Rating: 3.6666666666666665
Item: 9230-FLAWIL-Tankstellen BP, Estimated Rating: 3.6285
Item: 9200-GOSSAU-Tankstellen BP, Estimated Rating: 3.6285
Item: 9657-INTERMASSE-Hotels und Restaurants, Estimated Rating: 3.6285
Item: 7050-AROSA-Sportschulen, Estimated Rating: 3.6285
Item: 9400-RORSCHACHERBERG-Tankstellen AVIA, Estimated Rating: 3.6285
Item: 9422-STAAD-Tankstellen AVIA, Estimated Rating: 3.6285
Item: 3700-SPIEZ-Coop Mineralöl, Estimated Rating: 3.6285
```

Beispiel von Recommendations