

Machine Learning: Welche Android-App will weniger Permissions?

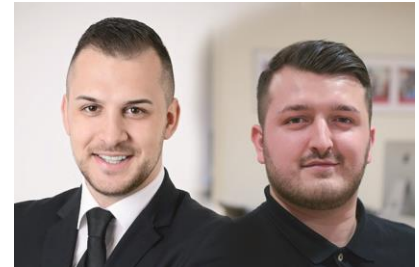
Der grösste App Store namens «Google Play Store» enthielt im Jahr 2020 über 2,6 Millionen Android-Apps. Diese unterscheiden sich in Zweck, Beliebtheit und Preis und benötigen unterschiedliche Berechtigungsstufen. Über den vorinstallierten Google Play Store können Drittanbieter-Apps mit Hilfe einer Stichwortsuche mit eingeschränkter Filter-Funktionalität gefunden werden. Im Google Play Store kann der Nutzer aus einer Liste von Apps auswählen, um detaillierte Informationen wie die Beschreibung, Screenshots und die erforderlichen Berechtigungen zu erhalten. Der Nutzer kann eine alternative App aus einer Liste ähnlicher Apps auswählen. Leider fehlt in dieser Liste eine Filter- und Sortierfunktion, um nach Apps zu suchen, die weniger Berechtigungen benötigen.

Deshalb beschäftigt sich die vorliegende Bachelorarbeit mit der Konzipierung einer Recommendation-Engine für alternative Android-Apps mit Filter- und Sortierfunktion, um die fehlenden Funktionen im Google Play Store auszugleichen. Die Recommendation-Engine bietet die Möglichkeit, nach Alternativen zu suchen, die weniger Berechtigungen erfordern. Sie verwendet die Metadaten der Apps aus dem Google Play Store in Kombination mit einem Topic-Modell, das aus den Beschreibungen der Apps generiert wurde.

Unsupervised Machine Learning Algorithmen wie Latent Dirichlet Allocation, Hierarchical Dirichlet Process und Non-Negative Matrix Factorization wurden auf bereinigte Teildatensätze angewendet, um den geeignetsten Algorithmus und die optimale Anzahl von Topics für den ganzen Datensatz zu evaluieren; die Empfehlungen der generierten Modelle wurden anhand von Googles Alternativ-Empfehlungen bewertet.

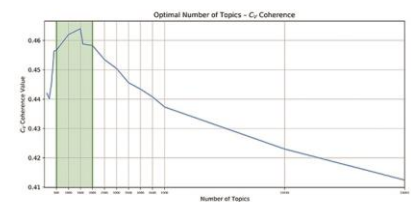
Zusätzlich wurde eine Web-Applikation entwickelt, die als Schnittstelle zur Recommendation-Engine dient. Die Web-Applikation ist containerisiert und besteht aus einem in JavaScript geschriebenen Frontend und einem in Python geschriebenen Backend. Das Frontend zeigt Android-Apps mit unseren Empfehlungen an und bietet Such-, Filter- und Sortierfunktionen. Die Backend-Services bedienen das Frontend mit Metadaten der App und den Topic Informationen.

Obwohl Google im Allgemeinen zufriedenstellende Empfehlungen anbietet, ist dessen Empfehlungsliste nicht optimal, um für Apps gefiltert zu werden, die weniger Berechtigungen erfordern. Es hat sich gezeigt, dass die vorgeschlagene Recommendation-Engine für diesen speziellen Anwendungsfall bessere Alternativen bietet.

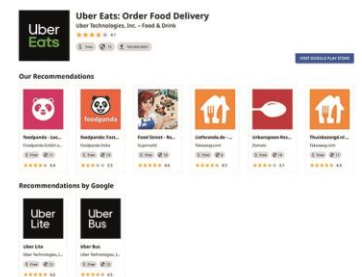


Diplomierende
Senad Jukic
Valmir Selmani

Dozent
Stephan Neuhaus



CV-Coherence von Topic Modellen, welche mit der Tomotopy Library und dem Algorithmus LDA mit einem Bereich von 100 bis 15'000 Topics erstellt wurden.



Vergleich zwischen unseren Empfehlungen und jenen von Google für die App "Uber Eats: Essenslieferung in deiner Nähe" (ID: com.ubercab.eats).