- Motivation: Ziel dieser Thesis ist der Entwurf und die Implementierung eines Framework, dass die Zusammenarbeit zwischen Entwickler (d.h. Analytics Engineers und Data Engineers) und Forscher (d.h. Data Scientists) fördert, um gemeinsam qualitätsgesicherte Code Artefakte bereitzustellen. Um den Machine Learning System Anti-Patterns Glue Code und Pipeline Jungles entgegenzuwirken, stellt das analyticstemplates Framework eine gemeinsame API in Palantir Foundry dar, die die Zusammenarbeit zwischen Entwicklern und Forschern ermöglicht.
- Methodik: Um die Anwendbarkeit auf datengetriebene Anwendungsfälle zu demonstrieren, implementieren wir eine wiederverwendbare Komponente als Palantir Foundry Code Workbook Template. Dieses Template löst ein Regressions Problem basierend auf einer gegeben Konfiguration und bildet somit einen Machine Learning Workflow ab. Der Proof-of-Concept basiert auf einem Anwendungsfall in der Preispolitik aus der Life Science Industrie, welcher die Rabatt Vorhersage, für einen Kunden der ein spezielles Produkt kaufen möchte, beabsichtigt. Um das Framework zu evaluieren wird eine Umfrage mit 12 Probanden durchgeführt.
- Ergebnis: Die Umfrage zeigt eine Durchschnittsnote von analytics-templates von 2.25 (1=sehr gut , 6=unzureichend). 91% der Befragten werden das analytics-templates Framework nutzen, wohingegen 100% zum Framework beitragen möchten. Darüberhinaus glauben 91% der Befragten, dass das analytics-templates die Auslieferungszeiten von Datenprodukten auf lange Sicht reduzieren wird.