

i4Ds18: Mit Machine Learning Immobilienpreise schätzen

Betreuer: [Manfred Vogel](#)
[Jonas Schwammberger](#)

| | | | |
|-----------------------|-----------------------|--------------------|--------------------|
| Arbeitsumfang: | P6 (360h pro Student) | Priorität 1 | Priorität 2 |
| Teamgrösse: | 2er Team | | P5 oder P6 |
| | | | 2er Team |

Sprachen: Deutsch

Ausgangslage

Vom Haus bis zum Parkplatz, wird für eine Schweizer Immobilie ein Käufer gesucht, so landet das entsprechende Inserat früher oder später auf den grossen Immobilien-Portalen wie [comparis.ch](#) oder [immoscout24.ch](#). Dort werden die Immobilien mit den wichtigsten Eigenschaften beschrieben: Anzahl Zimmer, Wohnfläche, Lage, Verkaufs-/Mietpreis, etc. Der Marktpreis einer Immobilie hängt von vielen Eigenschaften ab und ist deshalb schwierig festzulegen. Menschliche Immobilien-Schätzer greifen deshalb auf ihren grossen Erfahrungsschatz zurück und versuchen durch Vergleiche mit vielen anderen Objekten einen realistischen Marktpreis zu schätzen -- trotzdem weichen die so ermittelten Werte von verschiedenen Immobilien-Schätzer gerne einmal 15-20% oder sogar noch mehr voneinander ab.



Ziel der Arbeit

Diese Arbeit hat zwei Ziele:

- Als erstes soll ein (einfacher) Webseiten-Crawler entwickelt werden, um eine genügend grosse Menge von Immobilien-Inseraten zu crawlen, den Text zu extrahieren und in einer strukturierten Form abzulegen.
- Anschliessend sollen verschiedene Machine Learning Modell entwickelt, evaluiert und getestet werden mit dem Ziel, die tatsächlichen Verkaufs- resp. Mietpreise möglichst genau zu schätzen.

Problemstellung

Das Abschätzen von Immobilienpreisen ist ein Paradebeispiel für Lineare Regressions-Modelle. Aus "Gebäude-Features" wie Grösse, Lage, Baujahr etc. können mittels Machine Learning Immobilienpreise geschätzt werden. Machine Learning Ansätze profitieren von einer grossen Trainingsdatenmenge, bei denen die "Gebäude-Features" und der zu schätzende Preis bekannt sind.

Die Beschaffung von Trainingsdaten ist in vielen ML-Projekten aufwändig. In diesem Falle sind grosse Mengen von Immobiliendaten auf Online-Portalen vorhanden und können für die Entwicklung von Machine Learning Modellen verwendet werden.

Wird dieses Projekt als Bachelor-Thesis gewählt, erhöhen sich die Anforderungen sowohl an den geplanten Crawler als auch an die Machine Learning Modellen. Beispielsweise wird dann ein erweitertes Feature-Engineering erwartet und Lineare Regression-Modelle sollen durch weitere Ansätze wie Kernel Regression, Ensemble Learning, etc. ergänzt werden.

Technologien/Fachliche Schwerpunkte/Referenzen

- Für den Crawler: Programmiersprache Ihrer Wahl
- Machine Learning: Python/R/Weka/Matlab

Bemerkungen

Voraussetzung: Modul "Machine Learning" erfolgreich besucht.