Westfälische Hochschule

Fachbereich Informatik & Kommunikation
Show & Tell 2020

DataScout24 Datenanalyse eines Online-Kfz-Marktplatzes

Modul: Big Data Projekt (Wirtschaftsinformatik Master)

Team: Benjamin Ferfecki (WI), Enes Kara (WI), Matthias Janßen (WI)

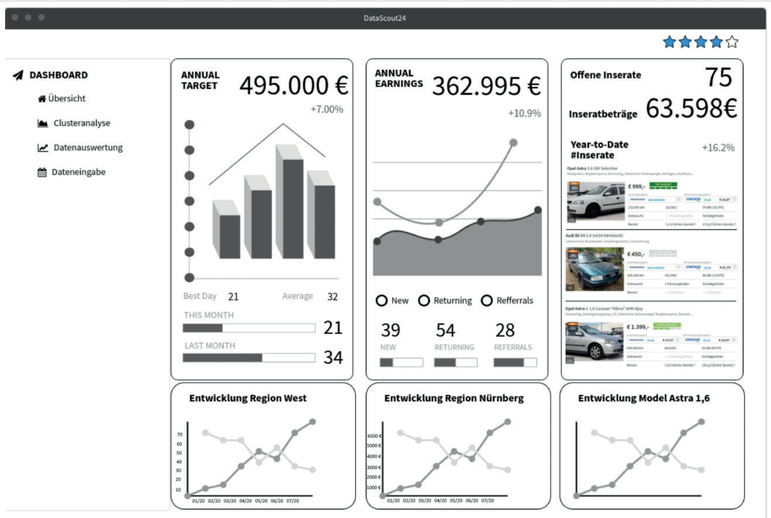
Problemstellung

- AutoScout24 bietet lediglich bei einzelnen Inseraten eine Einstufung des Inserats anhand des Preises im Vergleich zum errechneten Marktwert in fünf Kategorien von „Top Angebot“ bis „Teuer“. Weitere Einblicke in die Daten wie Entwicklung von Markt, Region, Marken oder einzelnen Modellen bzw. Antrieben gibt es Seitens AutoScout24 nicht.
- Hier gibt es Potenzial den Blick...
  a) auf weitere Parameter zu erweitern
  b) nicht nur Inserate zu betrachten, sondern auch die Entwicklung der Stadt/des Modells/der Marke/der Konkurrenz mit einzubeziehen, um den Verkäufern möglichst anschaulich mehr Insights in bereits vorhandene Daten zu geben und einen „Vorsprung durch Technik“ zu liefern und so langfristig durch bessere Marktkenntnis den Einkauf als auch Verkauf zu optimieren.

Idee und Konzept

Um eine Plattform für Verkäufer abbilden zu können, wird eine Datenbasis bestehend aus Inseraten benötigt. Da zur Zeit kein freier API-Zugang zu AutoScout24 verfügbar ist, wird dies über einen WebCrawler umgesetzt, welcher sämtliche Inserate automatisiert sammelt und abspeichern kann.

Diese Datenbasis wird dann genutzt um in der Web-Oberfläche Daten abzubildern, welche weitergehende Auswertungen beinhalten um auch den Markt, die Konkurrenz oder andere Parameter im Blick zu haben und so dem Verkäufer Vorteile auf dem Markt zu bieten. Alles „im Blick“ zu haben wird durch graphische Abbildung der Auswertungen gefördert.

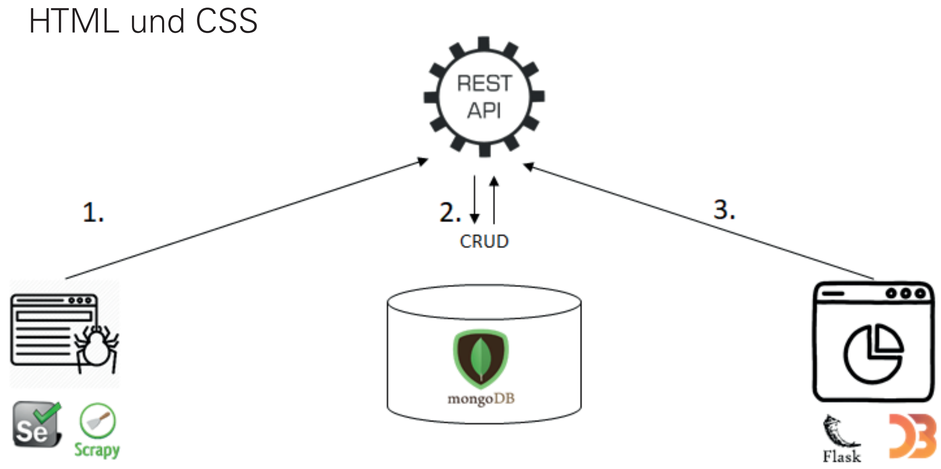


Mockup Web-Frontend Dashboard

Stand 20.01.2020 Web-Frontend Dashboard

Technische Umsetzung

- **Crawler** – WebCrawler zur Sammlung der Inseratsdaten (Marke, Modell, Ausstattung, Erstzulassung,…) unter Nutzung von Python, Selenium und Scrapy
- **RESTful API** – Zum zentralen Handling von CRUD-Abfragen auf die Datenbank unter Nutzung von Python und Flask (ein Micro-Web-Framework für Python)
- **MongoDB** – NoSQL Datenbank zur Speicherung der Daten
- **WEB UI** – WebInterface zur Darstellung der Daten in einem Frontend unter Nutzung von Python Flask, Jinja2, JavaScript, D3.js, HTML und CSS



Team

Benjamin.Ferfecki@studmail.w-hs.de
Matthias.Janssen@studmail.w-hs.de
Enes.Kara@studmail.w-hs.de

Betreuung

Prof. Dr. Henning Ahlf
Fachgebiet: Wirtschaftsinformatik

Kolja Dunkel
Fachgebiet: Wirtschaftsinformatik