Used Car Appraisal

Thinkful Final Capstone by Luke Elliott

Data and Objective

- The Data
    - From on Kaggle.com
    - Scraped from eBay Kleinanzeigen
    - 370,000 Used car listings
        - Price, year, mileage, gearbox, etc.
- Objective
    - Predict price of car based on other features
    - Scrape current website and assess value of cars

Data Exploration

- Dealing with Missing Data
    - Exploration turned up much missing data.
    - Imputed information based on other features
        - Horsepower from median horsepower for that model of car
        - Model from the brand and name of the listing
- Assessed usefulness of each feature
    - Checked to see impact of individual features on the outcome variable.
    - Deleted features with low/no variance or little predictive value.

Predictive Models

- Sampled dataset and prepared it for modeling
    - The dataset was too big for my computer to handle.
    - Created dummy features out of categorical variables
- Tried various models from scikit learn
    - K Nearest Neighbors was slow and inaccurate.
    - Random Forest and Gradient boosting did well with 'r2' scores of roughly .82 and ran relatively quickly.
- Implemented a neural network with TensorFlow
    - Tuned it with GridSearchCV
    - Proved less effective than the scikit learn models

Web Scraper

- Used Scrapy to get a dataset similar to the one I started with.
    - Built translation into the scraper using a dictionary
- Cleaned it up to make it match the other dataset.
    - Fixing data types
    - Changing spelling conventions
    - Accounted for new terms on eBay Kleinanzeigen

Results

- Estimator with .82 r2 score on test set.
- A product for quickly finding cheap cars on eBay Kleinanzeigen

	assessed value	margin	name	price
0	10994.86	994.86	bmw 525xi m paket xenon navi	10000.0
1	3760.50	-1019.50	nissan micra 1.2 edition 25 jahre 1 jahr garan...	4780.0
2	66238.44	-53261.56	porsche 991 carrera 4s cabriolet sportchrono s...	119500.0
3	859.73	-140.27	volkswagen golf 1.6	1000.0
4	4853.90	403.90	mini cabrio sonderausstattung mit tüv leder s...	4450.0
5	13750.96	1260.96	volkswagen passat cc 2.0 tdi bluemotion techno...	12490.0
6	6896.97	-4083.03	smart fortwo coupe passion/cool&audio/jbl/led/...	10980.0
7	7215.12	-5744.88	smart fortwo coupe prime/twinamic/cool&audio/l...	12960.0
8	3135.62	35.62	opel meriva 1.8 16v cosmo mit gasanlage	3100.0
9	4425.12	2625.12	audi a4 2,5 tdi 6gang 179ps quattro tüv 12.19 ...	1800.0
11	3404.68	1904.68	bmw 525d automatik exklusive fast voll tausch	1500.0
12	7197.73	-5372.27	smart fortwo coupe passion/twinamic/navi/led/s...	12570.0
14	3427.74	1426.74	toyota yaris basis, neue ganzjahresreifen, tüv...	2001.0
15	3466.75	-2233.25	1996 ford (usa) ranger pickup	5700.0
16	5685.09	3886.09	vw golf 5. 1,9 bls bluemotion 5 gang. motor ge...	1799.0
17	5031.77	-968.23	verkauf mein bmw 320	6000.0
18	28936.20	5036.20	volkswagen touareg4.2v8r-line+exclusive wirkli...	23900.0
20	8568.09	569.09	a4 avant 1.8 tsi s-line, vollausstattung schec...	7999.0
22	11262.54	-537.46	bmw 130i m-paket,m-performan	11800.0