Ketul Padariya

Machine Learning Engineer

04 August 1994

Ahmedabad

+91 90336 99226

www.github.com/KETULPADARIYA

ketulpadariya79@gmail.com

Skills

- Deep Learning
- Data Augmentation
- Data Visualisation
- PySpark
- Gensim
- Python
- NLTK
- Spacy
- Tensorflow
- Sklearn
- NLP
- Python
- Testing
- SQL and Elastic Search

(*)[The skill scale is from 0 (Fundamental Awareness) to 6 (Expert).]

Summary

Experience

2019 OCT — Machine Learning Engineer — VALENS DATALABS - Ahmedabad

In the latest project, 3 million products corpus were created and cleaned using NLTK and the Spacy Library, followed by a handmade CNN model to predict 2400 classes. The corpus was further divided by the 2400 class model. Then for each class, the entity extraction model was used to extract entity from the document.

To obtain identical documents from corpus, first the documents was vectorized by the Fasttext model, the TFIDF model, and the LSI model, then the same matrix class of the Gansim library was used for faster extraction of the similar documents.
Basket analysis was used to obtain product's associative products.

In the first project, I developed a project for entity tagging in real word random text data and data augmentation.
After that, the Stanford and Space Entity tagging models were used to extract noun phrase from the text. The predictive power of the entity was further enhanced by the statistical model based on the part of the speech tag.
Various statistical tests were used to compare the performance of the models
Library used : tensorflow, pyspark, Sklearn, Spacy, Nltk, scipy, pandas, numpy, gensim, elasticsearch, pymongo, kafka, SQLite, matplotlib, streamlit, seaborn, flask

Internships

2019 AUG — Power BI Developer — IN KEY SOLUTION -Surat

Collaborated with data engineers and operation team to implement the ETL process, wrote and optimized SQL queries to perform data extraction to fit the analytical requirements.
Explored and analyzed the customer specific features by using dashboards in PowerBI . Technology Use: PowerBI,SQL,excel.

2019 JAN — Web Scrapper and Data Analyst — WEB SECURE -Surat

Used the BeautifulSoup and Scrapy Python Library to extract desired data from a given URL.
Experimented and built predictive models including ensemble methods such as Gradient boosting trees and Neural Network by Keras to predict Sales amount.
Worked on data cleaning and ensured data quality, consistency, integrity using Pandas, NumPy.
Explored and analyzed the customer specific features by using Matplotlib, Seaborn in Python and dashboards in Tableau.
Technology or Library used: BeautifulSoup, Scrapy, Pandas, Numpy, PoweBi, Matplotlib, Django, Seaborn, Tableau, SQLlit

Education

2018 — PEC UNIVERSITY OF TECHNOLOGY, Chandigarh — M.tech
with the average of 79 percentage

2016 — SHANTILAL SHAH ENGINEERING COLLEGE, Bhavnagar — B.tech
with the average of 75 percentage

Ketul Padariya

Machine Learning Engineer

04 August 1994

Ahmedabad

+91 90336 99226

www.github.com/KETULPADARIYA

ketulpadariya79@gmail.com

Skills

- Deep Learning
- Data Augmentation
- Data Visualisation
- PySpark
- Gensim
- Python
- NLTK
- Spacy
- Tensorflow
- Sklearn
- NLP
- Python
- Testing
- SQL and Elastic Search

(*)[The skill scale is from 0 (Fundamental Awareness) to 6 (Expert).]

Other information

In the last 3 years, I have actively learn programming and math for AI from over 20 courses and over 10 books. Below are some of the courses I have completed.

Courses

- Learn C++ Programming- Beginner to Advance- Deep Dive in C++ (Udemy)
- Mastering Data Structure and Algorithm using C and C++ (Udemy)
- Machine Learning (Andrew Ng : Stanford, Udemy)
- Deep Learning (nptel, Udemy)
- Natural Language Understanding (Stanford)
- Natural Language Processing (Nptel)
- Complete Guide to Elastic Search (Udemy)
- Pandas, Probability-Statistics, Shell, PyTest, StracutruingCode (DataCamp)
- Microsoft Power BI (Udemy)
- Probability - The Science of Uncertainty and Data (Mit-edx)
- Reddis boot camp for Beginners Get Hands On with Redis (Udemy)
- Deep-Dive into Python - OOC (Udemy)
- Design-Patterns in Python (Udemy)
- Excel and advance Excel (Microsoft-edx)
- etc.

Books

- Python for Data Analysis : Data Wrangling with Pandas,Numpy by Wes McKinney
- Hands on machine Learning With Scikit - Learn ,Keras and TensorFlow-2 by Aurelien Geron
- Introduction to Linear Algebra by Gilbert Strang
- Neural Network and Deep Learning -Michael Nielsen
- Natural Language Processing With Python by Steven Bird
- Data Structure and Algorithm in python by Roberto Tamassia
- Head First SQL: Your Brain on SQL – A Learner's Guide by Lynn Beighley
- etc