Himanshu Kriplani

linkedin.com/in/himanshu-kriplani/ | github.com/Himanshk96/ | kriplani@usc.edu | Los Angeles, CA | +1 (213) 378 9578

EDUCATION

University of Southern California: **August 2019-May 2021 (Expected)**
Master's in computer science with specialization in Data Science **GPA:3.2/4.0**
Courses: Machine learning, Database Systems, Data Mining, Applied NLP

Ganpat University: **August 2015- April 2019**
Bachelor of technology, Computer science with specialization in Big Data and Analytics. **GPA:9.36/10**
Courses: Big data and analytics, Machine learning, Data warehousing and mining, Business Intelligence, Software Engineering

TECHNICAL SKILLS

- **Programming languages:** Python, Scala, R, C++, Java
- **Databases:** MongoDB, Cassandra, Hbase, and SQL
- **Big Data Tech:** MapReduce, Spark, Kafka, MLlib, Hive, Pig
- **Miscellaneous:** Numpy, Scikit-learn, AWS(S3 and EC2), Keras, GCP, Git

EXPERIENCE

Graduate Research Assistant, Minerva Lab,USC (Los Angeles, US) **December 2019-Present**

- Analyzing startup pitches and developing an automated pipeline for inference and statistics.
- Implementing LDA on the dataset to extract topics and do topic modelling on the dataset.

Data Scientist Intern, Embibe (Bengaluru, India) **December 2018-May 2019**

- Collaborated with smart tagging team and implemented image classification based on CNN and Resnet architectures.
- Implemented text classification using ULMFiT improving the accuracy by 15%.
- Automated scripts for data extraction (SQL) and cleaning saving time by almost 50%.
- Wrote NLP utility library adopted by whole data science team.

Machine Learning Intern, Strawberry Labs (Ahmedabad, India) **June 2017-July 2017**

- Worked on text matching model using tf-idf for chatbot reducing support wait time of customers by 20%.
- Designed an android app for its interface.

PROJECTS

- **Recommendation engine on Implicit Dataset:**
  - Performed exploratory data analysis on data with pandas and matplotlib.
  - Used Alternating least square and LightFM on dataset to create Recommendations
- **Text Classification using ULMFiT:**
  - Used Spacy to pre-process the dataset.
  - Used ULMFiT pre-trained model to create a text classifier on the dataset.
- **Market Basket analysis and Recommendation using Spark**
  - Programmed several algorithms like Apriori, LSH using Spark.
  - Used MLlib for model based collaborative filtering. Additionally, build item based CF models for recommendation.
- **OneGate (A social network for developers):**
  - Merged all the KPI's of Stackoverflow and github into a single profile-based web-app build using React and was awarded best in regional Facebook's HackDay.
- **LifeBlood (A blood donor's app):**
  - Implemented a PHP webapp with Mysql DB to connect Donors and Seekers.
  - Improved search algorithm and could get back most relevant donor when searched.
- **Sampler in R:**
  - A shiny web-app to sample (to select a small amount of data from huge dataset) data based on different sampling techniques such as SMOTE, ROSE, Stratified Sampling, and others.

PUBLICATIONS (Research Papers and Blogs)

- **Author** Prediction of arteriovenous nicking using aggregate residual network. International Conference on Computational Intelligence in Data mining (2018) (Link)
- **Blog** Recommendation engine using Alternating Least Square method (Link)
- **Blog** Understanding language modelling and ULMFiT (Link)

LEADERSHIP AND EXTRACURRICULAR ACTIVITIES

- **Coordinated Science and Tech Club** by facilitating student activities at Ganpat University (2016-2019).
- **Volunteered for teaching and mentoring** government school students.
- **Director General Student Achievement Award (2018) [Ganpat University]**

CERTIFICATIONS

- Data Scientist 2017v2 by IBM (2018)
- Apache Spark Essential Training (Linkedin Learning)
- Statistics Foundations (Linkedin Learning)