

# SHRADDHA KULKARNI

(213) 477-3598 | [shraddhakul@fb.com](mailto:shraddhakul@fb.com) | <http://shraddhak.blog> | [linkedin.com/in/shraddha-k27/](https://www.linkedin.com/in/shraddha-k27/)

## EDUCATION

**University of Southern California**, Los Angeles, CA  
M.S. in Data Informatics | Aug'18 – May'20 | 3.5/4.0

**LDRP-ITR, Kadi Sarva Vishwavidyalaya**, Gujarat, India  
B.E., Computer Engineering | Aug'14 – May'18 | 8.8/10.0

## SKILLS

**Languages & Dev Tools:** Python, Java, C++, PHP, Unix, HTML5, CSS3, JS, ReactJS, Bootstrap, JSON, Android Studio

**Database:** XML, JSON, SQL, NoSQL, Big Query, Firebase, MongoDB, Dynamo DB

**Analytics & Modelling Tools:** Tensorflow, Scikit-learn, Spark RDD, SparkML, Spark Graphx, Apache Beam, Tableau, Google Analytics, Data Studio, JMP, Visio, Stackdriver Monitoring, Cloud Dataflow, Google Dataproc, Cloud Pub/Sub

**Major Coursework:** Probability and Statistics, Data Mining, Machine Learning for Big Data, Data Visualization

**Certifications:** Google Certified Professional Data Engineer, Managing Big Data with MySQL [Duke University], Building Resilient Systems on GCP [Google Cloud Certified]

## WORK EXPERIENCE

**Data Engineer**, *Facebook Inc., Menlo Park, CA*

June 22, 2020 – Present

- Part of the Facebook App - Connection Integrity team.
- So far worked on building efficient pipelines and experimentation modules.
- Collaborated with Data Scientists to come up with modelling solutions to track the team's impact on Facebook ecosystem.

**Full Stack Developer Intern**, *Los Angeles Times Inc., Los Angeles, CA*

Jan 23, 2020 – Apr 21, 2020

- Worked simultaneously on two projects using diverse set of tools like React, Next.js, Graphene CMS, Node.js, SQLite, Figma.
- Built a Restaurants Application, "Where to Eat?", to aggregate and display the curated list of Food Places/Restaurants in Los Angeles County. The first version was built in a matter of 4 weeks, which included User Testing, Designing, Database Setup, Data Formatting, Plugin with existing CMS System, and Application Building.
- The second one was an NLP Product for Document Analysis. It included dashboards for Information Visualization.
- Developed a Front-end for an Associator Product which helped the users gain insights by suggesting the related/associated words given an input. All the edge cases were well-handled along with a cleverly designed User-interface. One of the dashboards included a Heatmap which showed the frequency distribution of the inputted word page-level wise.

**Machine Learning Engineer Intern**, *Pluto7 Consulting Inc., San Jose, CA*

May 28, 2019 – Aug 23, 2019

- Applied advanced analytics to understand Correlations and patterns on real-time data generated during the production of beverages at the AbinBev factory, to select 10 most important features out of 60+ features.
- Formulated various Clustering strategies to predict the optimal setpoint values for the production cycles which ultimately led to optimized Cycle time and effective gas usage reduction.
- Cleaned & Analysed historical sales data of Gelson's Market, with 5K+ products, 25 GB+ of raw data to predict product sales.
- Leveraged GCP Platform to create ETL Pipeline, deploy models, and get the online and batch predictions.
- Followed the rigorous AI/ML Cloud Platform Coding standards for ML Specialization Project, which is to be evaluated by the folks at Google to certify the skillset of ML/Data Engineers at the firm.

**Data Science Intern**, *Tatvic Analytics, India*

Dec 15, 2017 – May 06, 2018

- Built a Remarketing Deep Learning model which resulted in reduced CPL & 1.5x Increased ROI on PPC Marketing Campaigns
- Developed a valuable SEO tool that suggested the keywords to bid on, based on the Search and Volume trends of Client's most frequent Keywords. The script can analyse and store historical data to the Google Cloud.
- Performed Market Research using Google and FB Analytics data to provide insights regarding Market trends, competitors and promotional campaigns for customers using Interactive UI Dashboards.

## PROJECTS

**AI Research Fellowship Project** (*Among the 10 scholars around the globe, to work on an independent project*)

- Researched AI Algorithms for Traveling Salesman Problem to output shortest path covering 500+ stops spread across NYC.
- Designed ETL Pipelines to transform 21M+ rows using BigQuery, Dataflow Runner, and Apache Beam scripts.
- Deployed multiple Tensorflow & Scikit-learn Prediction models using Neural Network & Decision Trees Algorithms on Cloud.
- Used Spark Graphx, SparkML, BigQuery and Pandas Manipulation techniques to analyse the trends in data for Time Series and Clustering Techniques.

**Predictive Classifiers for Automated Detection of Glaucoma** (*Research paper in the American Journal of Ophthalmology*)

- Pre-processed the Anterior Segment OCT Images to adjust its dimensions and segregate the individual portions.
- Analysed the images to select appropriate CNN and Transfer Learning Models for detection of the Eye disease.

**Camfit - Camera Compatibility**

- When buying a camera and/or its accessories, the user needs to do a lot of back-and-forth online searches to decide which products are compatible. Furthermore, this question of compatibility can be multi-layered. Currently, no other website produces a thorough list of compatible products, therefore making a purchase is always a challenge.
- To this end, we constructed a knowledge graph and developed a web interface. Our app allows users to do layered faceted filtering and quickly generate lists of compatible products.

#### **Machine Learning powered Cloud based Search App**

- Developed a search web application using HTML, CSS, JS, and Django backend Interface. Scraped data through online sources and stored on MongoDB Server. Hosted the application on AWS Instance.

#### **Implementation of Machine Learning / Deep Learning Algorithms from Scratch**

- Implemented Supervised, Semi-Supervised, Unsupervised Learning on Diagnostic Dataset for classification problem.
- Reduced 170+ columns using Feature Selection techniques and Handled missing and Class imbalance problem.
- Applied Hyperparameter Tuning, CV and Monte-Carlo Simulations during the implementation to increase the confidence in the obtained results.

#### **Apache Spark-based Collaborative Filtering Recommendation Engine**

- Built a Recommendation engine using concepts of Pearson Correlation, Minhashing, and Least Sensitive Hashing with Spark RDD and Spark ML APIs. The running time was 40 seconds for more than 1.5 GB of local data processing.

#### **Flask-based Twitter Analytics Dashboard**

- Created a dynamic Dashboard in Flask using Spark RDD (Pre-processing), Chart.js, and Twitter APIs to display the top 10 Hashtags used in the past 2 seconds around the globe on a particular Topic.

#### **Image Recognition Android Application for Blind (*Among Top 5 most innovative Departmental Projects*)**

- Developed an app targeted towards Blind community applying Google OCR and Microsoft CV APIs for Object, Emotion, and Character Recognition, with in-build Voice support for convenience.

## **PUBLICATIONS**

---

### **Deep Learning Classifier for Automated Detection of Gonioscopic Angle Closure Based on Anterior Segment OCT Images**

*Research Paper* – Published in American Journal of Ophthalmology, Presented at the World Glaucoma Conference, Australia

### **Cloud Computing in India: Challenges, Progress and Future** *IJRASET [ISSN: 2321-9653, Volume 4 Issue XII, Dec 2016]*

*Research Paper* – Inspected India's current scenario in the field of Cloud Computing, collected & integrated relevant information.

### **Study of Modern Cryptographic Algorithms** *IJARCS [ISSN No. 0976-5697, Volume 8, No. 3, Mar 2017 – Apr 2017]*

*Review Paper* – Assessed various cryptographic algorithms and compared their efficiencies in MATLAB.

## **CERTIFICATIONS & LEADERSHIPS**

---

### **Certifications:**

- Google Certified Professional Data Engineer,
- Managing Big Data with MySQL [Duke University],
- Building Resilient Systems on GCP [Google Cloud Certified]

### **Leadership & Volunteering:**

- Publicity Head, IEEE, LDRP-ITR, Student Chapter, Gujarat, India. [Feb 2017 – May 2018],
- Volunteer at PyData LA and IDEAS Conference held at the LA Convention Centre,
- Volunteer, Art of Living Organization, India. [2010 - Present]