

# ABHISHEK JANGALWA

1230 W 29 St, Los Angeles, CA 90007 | +1 213-210-1939

[jangalwa@usc.edu](mailto:jangalwa@usc.edu) | <http://linkedin.com/in/jangalwa> | <http://jangalwa.me> | <https://github.com/jangalwa>

## EDUCATION

**University of Southern California**, Los Angeles, CA

August 2017 - May 2019

M.S. Computer Science- Data Informatics, GPA - 3.35

Courses: Machine learning, Data Management, Data Mining, Knowledge Graphs, Information Retrieval, Capstone Project

**Rajiv Gandhi Proudyogiki Vishwavidyalaya**, Bhopal, India

August 2006 - June 2010

Bachelor of Engineering, Computer Science, GPA - 4

Courses: Analysis and design of algorithms, Data structures, Object oriented programming, Database

## TECHNICAL SKILLS

- **Programming Languages:** Scala, Python, PHP, SQL, JavaScript, HTML, CSS
- **Databases:** MongoDB, MySQL, Firebase, Oracle, HBase, Cassandra, DynamoDB, Hive
- **Libraries & Frameworks:** Spark, Hadoop, JQuery, Ajax, Flume, Scikit Learn, Flask, Matplotlib, Tensorflow, RDF
- **Systems & Tools:** Amazon AWS, Unix, Linux, Windows, Jupyter, Git, Six Sigma, MapReduce

## WORK EXPERIENCE

**Research Engineer Intern (Data Science)**

January 2019 – Present

**Evidation Health Inc., Santa Barbara, CA**

- Developing data lake for an internal project that will help serving data on demand to other projects.
- Scrapping open source data sources to fetch data for lake, clean and store it in Parquet format. This part is run as a cron job on a Linux server on daily and weekly basis.
- Created a Flask based REST API that serves the data, also a Python API that wraps up the end points and queries the lake for easier access to the data.

**Machine Learning Pipeline Developer (Graduate Research Assistant)**

November 2017 – January 2019

**USC Information Sciences Institute, Marina del Rey, CA**

- Developed ML pipeline using Spark for TILES (Tracking Individual Performance with Sensors) study that enables Data Scientists to extract features from raw sensor data and start creating models within minutes, also providing means for model evaluation, cross validation and hyper-parameter optimization.
- Wrote and re-factored existing Python code for managing, cleaning and transforming data, collected from 6 sensors and various surveys. Improved speed of legacy data aggregation modules up to 8000%.
- Constructed custom PySpark modules for time series data analysis of 20 TB data.
- Visualized data using Matplotlib, seaborn and d3 to help researchers better understand data analysis results, also maintained project website <http://sail.usc.edu/tiles>.

**Full Stack Developer, Founder Director**

January 2010 - April 2017

**Paritosh Software Private Limited, Mandsaur, Madhya Pradesh, India**

- Spearheaded day-to-day business operations with an emphasis on product development and outreach.
- Identified and suggested new technologies and tools to enhance team productivity by 70 percent.
- Built and maintained client relationships, resulting in revenue and profitability growth.
- Trained and mentored over 100 engineering students, helped begin a career in web development.

## PROJECTS:

**Academic Projects at USC**

- **Recommender systems:** Implemented user based and item-based recommendation systems using Jaccard and Cosine similarity with Scala and Spark. Also implemented model-based recommendation system using Spark MLlib on Amazon Review Dataset.
- **Software for Prediction of Liver Transplant recipient using Machine Learning:** Created software with GUI, based on dataset provided by UNOS and Keck Medical Center of USC using Python, Flask, Scikit Learn, Tensorflow and Keras. Also utilized Six Sigma principles of project management.
- **Knowledge graph (About Books):** Built knowledge graph about books. Utilized Scrapy, BeautifulSoup for website crawling and Snorkel, Spacy for information extraction. Performed entity linking and RDF creation.
- **Neural network:** Achieved 89.1% accuracy with multi-class image classifier, using feed forward neural networks in Python by applying back-propagation algorithm without using any libraries.
- **Website for Adventure Gurus:** Designed website using Flask (Python web framework) and MongoDB, deployed on Amazon EC2 as a volunteer and extracurricular.

**Websites and LAN Based Softwares**

January 2010 – April 2017

- Drafted fully-responsive dynamic website from scratch and programmed solutions using Linux-Apache-MySQL-PHP (LAMP) for billing, inventory management, online ticketing and other commercial use cases.
- Conceptualized and created software and web solutions for more than 50 government and private clients including Administrations, Educational Institutes, Manufacturers, Hospitals and Agricultural Produce Market Committees.