

# KETAKI LOLAGE

(213) 551-5690 | [klolage@usc.edu](mailto:klolage@usc.edu) | Los Angeles, CA | <https://github.com/ketakilolage> | <https://www.linkedin.com/in/klolage234/>

## Professional Experience

### Software Developer Intern, nēdl

May 2022–July 2022

- Saved \$30,000 in transcribing 1000 podcasts by replacing AWS Transcribe with Mozilla DeepSpeech in Python
- Saved 58 hours of manual labor by scraping Wikipedia, Audials, and radio-browser for radio stations
- Replaced the trending words module; exposed the Node.js backend APIs; documented architectural revisions
- Managed the old MongoDB database; migrated users to the new PostgreSQL database

### Junior Consultant (Developer), TIBCO Software

Sep 2018–Aug 2020

- Designed and developed RESTful SaaS enterprise integration solutions using TIBCO tools, Java, and SQL
- Spearheaded a TIBCO Cloud Integration project to deploy microservices enabling data pipelining between Salesforce, Amazon Athena, and TIBCO EBX, on a Kubernetes cluster running on AWS EC2

## Skills

**Languages:** Python, Java, C/C++, JavaScript, Shell scripting

**Databases:** Oracle SQL, MySQL, MongoDB

**Web Development:** HTML, CSS, Node.js

**OS:** Linux Ubuntu, Windows 7+

**Other:** Jupyter Notebook, VS Code, Eclipse IDE, Docker, Git, TIBCO technologies (Certified)

## Projects

### Comparative Analysis of Genetic Data for Anomaly Detection (Team of 4)

- Implemented a hierarchical framework to calculate chances of a subject being afflicted by genetic diseases
- Reduced computation time e.g. computation against dengue sequence finished in 1.2 hours on the hierarchical setup as opposed to 24 hours on the PARAM SHAVAK Supercomputer
- Tools: Python, MongoDB, PHP, HTML, CSS

### Time Series Analytics

- Fitted logistic regression models to time-domain features extracted from time series data of human activities
- Performed cross-validated recursive feature elimination on differently shaped splits of the dataset
- Corrected marginal imbalance using SMOTE; achieved reliable coefficients
- Tools: Python, Jupyter Notebook

### Predicting Violent Crime in Communities

- Analysed the effect of 122 socio-economic factors on violent crime per capita in a community using 5 methods: ordinary least squares linear regression, ridge regression, LASSO, PCR, and boosted decision tree
- Achieved 98.27% test accuracy using the boosted decision tree
- Tools: Python, Jupyter Notebook

### Blood Pressure Estimator

- A UI interface to a linear regression model that predicts blood pressure based on age and weight
- Tools: Flask, HTML, CSS

## Education

### University of Southern California

Los Angeles, CA

Master of Science in Computer Science

Aug 2021–May 2023

### University of Pune

Pune, India

Bachelor of Engineering in Computer Engineering

Jul 2014–May 2018

## Publications

- Co-author of "Comparative Analysis of Genetic Data for Anomaly Detection", RACE Mar 2018
- Co-author of "Genetic Sequence Alignment: A Comparative Study of Methods", IEEEExplore Feb 2018