Jeevan Vankayala
611 Lead Ave SW, Albuquerque, New Mexico.
E-mail: jeevansai2@gmail.com Ph. 505-307-8697 Pin: 87102.

ACADEMIC QUALIFICATION

- **Masters in Computer Sciences** ***August '14 – Till date***
  University of New Mexico, Albuquerque. GPA: 3.36
- **Bachelor of Technology in Computer Science and Engineering** ***July '07 – April '11***
  Amrita School of Engineering, Amrita Vishwa Vidyapeetham University, Coimbatore. CGPA: 7.41*/10*

PROFESSIONAL EXPERIENCE

Python Developer,* Bureau of Business and Economic Research, UNM** ***Jan '15 – Till date

- Automated the process of data cleaning for Geo-Spatial and Population Studies. Developed applications in Python to connect to MySQL database using MySQL connector, reading from excel sheet and perform regular expression matching to implement and optimize the data transformation processes. Python, Pandas, Maria DB, MySQL.

Senior Systems Engineer,* Infosys Limited, Hyderabad, Telangana, India** ***July '11 - June '14

AOSS (Apple Online Sales and Services) Reporting for Apple Inc.

- For 4+ years I demonstrated strong skills in data warehousing to design complex logic's in Teradata PL/SQL, BTEQ, Fastload, Multiload, Fast Export and Informatica mappings to load data aggregations into Teradata for reporting, dash boarding and ad-hoc analysis.
- Wrote ad-hoc SQL queries to perform data analysis, troubleshooting and performed performance tuning.
- Developed Python and UNIX applications that automated ETL pipelines, data loadings and data validations along with various re-usable tools that find data mismatches and triggers an automated email and a P1 ticket to support.
- Cleared Teradata 12 Basics exam TE0-121 in March '13 and became **Teradata 12 Certified Professional**.

INTERNSHIP

Infosys Limited, Mysore, Karnataka, India ***Feb '11 – June '11***

- Implement steganography algorithm such as Least Significant Bit (LSB) algorithm to embed encrypted data into MPEG media files so that it can be securely transferred to an intended recipient. JavaFX has been used to develop interactive user interface.

ACADEMIC PROJECTS (www.github.com/jeevan4)

Recommender System on Million Songs Dataset: ***Nov '15***

- Implemented Collaborative Filter Recommendation system on 1M song dataset that contains user-song play counts. Implemented the project in Apache Spark to build the recommendation model after doing appropriate normalizations. Achieved RMSE of 0.082.
- Improved the above model into Content Based Filtering to extract similar tracks from MongoDB using K-means clustering. Precision and Recall has been implemented to evaluate the model.

Ticketing System for Office of Institutional Analytics: ***Nov '15***

- Developed a ticketing web app using Ruby on Rails, HTML5, JavaScript, CSS and MYSQL to automate the data approval process. Automated emails were sent using Mandrill API once a ticket is requested or serviced. The application has been deployed into Heroku and currently used by the department. (https://bitbucket.org/jeevan4/webarch-final) (https://oiaticket.herokuapp.com)

Sentiment Analysis on IMDB and Yelp reviews: ***Oct '15***

- Obtained 1.5M Yelp reviews from Yelp's Academic dataset, scraped 50k IMDB reviews and labelled as +ve and –ve based on ratings. Pre-processed data by removing html tags, junk symbols, stop words and performed stemming.
- Implemented Hadoop MapReduce for IMDB and Apache Spark for Yelp data to extract features as bag of words and bi-grams to represent the reviews. Used Random Forest machine learning technique to identify the polarity of the test dataset. Achieved accuracies of 85% and 72.5% for IMDB and Yelp respectively. Implemented in Python using Hadoop Steaming APIs, Apache Spark, Hive, Nltk, Beautifulsoup, sklearn, Numpy.

Application of Machine Learning to classify music files and newsgroup document: ***May '15***

- Implemented multinomial logistic regression with gradient descent technique that can classify a music file based on the genre. The project has been implemented in Python, Numpy, Scikit learn and achieved accuracy of 51.33%
- Goal of this project is to predict the give classic newsgroup a given document was posted to. Used Naive Bayes technique to classify documents and achieved accuracy of 78.5%. The project has been implemented in Python, Numpy.

Deal Hunt (Android App): ***May '15***

- An android app that aggregates deals from various retailers through their RSS feeds and show's all of them in one place. Deals can be saved or removed from wish list and can also be browsed offline. App pushes a notification if a saved deal is about to expire, so that you will never miss a deal. Android SDK, XML parser.

TECHNICAL SKILLS

- Python, Hadoop MapReduce, Apache Spark, Hive, Pig, Sqoop, scikit-learn, Numpy, Pandas, C, C++, Java, PHP, UNIX scripting, HTML5, CSS, JavaScript, Ruby on Rails, Android SDK, Informatica, Business Objects, Cascade Servers, and GIT.
- Databases: Teradata, MongoDB, Postgres, Oracle, SQL Server, MySQL, Maria DB.