Jose Alex

Vachaparambil House,
Bangalore- 29

Ph. 8050001023
GitHub - https://josevacha.github.io/online/
LinkedIn - www.linkedin.com/in/josealex
Email - mail2josephalex@gmail.com

INTRODUCTION

Confident and articulate individual with 1.5 of years of experience in supporting machine-learning projects. Proven hand on experience in Content Engineering, Natural Language Processing (NLP), Statistical Modelling, Web Scrapping and Data Visualization. Experience in working closely with clients to understand their business needs and providing end-to-end solution, which results in achieving their business goals.

ANALYTICAL STACK

Scientific/Statistical Computing:

- Python.

Libraries:

- NumPy and Pandas
- Scikit-learn
- Seaborn
- Matplot

Tool & Databases:

- SQL
- Oracle
- Mongo DB
- Apache Spark
- MS Excel

EXPERIENCE

Support Engineer | Meltwater Software Pvt Ltd
May 16 – Jun'17

Project Name: Meltwater Feeds (MWF)
Project description: Introduced and lead a team of four, in development of a new web-scrapping robot using Python and Mongo-DB. An intuitive tool with dynamic data extraction capabilities. A project well correlated with the company's goals, which made a significant impact on the company financials.

MODULE DESCRIPTION:

- **Web Scraping** – The data extraction module was completely designed and developed in python using beautiful soup library. Data is scrapped from specific HTML tags periodically and were stored in Mongo-DB. The tool was capable of scarping data even from a JavaScript-enabled-websites.

- **Basefarm Server –** Hosted in Basefarm and setups a CRON job to fetch data periodically. The data is sent to content service API and then to the enrichment pipeline for language detection, text summarization, and sentimental analysis. The enriched data is index by elastic search and stored in RIAK database.

DATA SCIENCE STACK

Algorithms Used:

- Linear Regression
- Logistic Regression
- Decision Trees
- Support Vector Machines (SVM)
- Neural Networks with Gaussian Kernel
- K-Means
- Data Compression using PCA

Programming Languages:

C, C++, Java and Python

EDUCATION

Master of Computer Application
Christ University
Bangalore - 2016.

Bachelors in Computer Application
Christ University
Bangalore - 2013.

Pre-University College (PUC)
Vijaynagar College
Bellary – 2010

CERTIFICATIONS

Machine Learning – IIM
Data Centre and Analytics Lab (DCAL)
Indian Institute of Management
Bangalore.

Machine Learning by
Stanford University on Coursera

Deep Learning – Andrew NG
(Currently pursuing)

Data Mining, Predictive Analytics, and Big Data
School of Professional Studies
New York University (NYU)
(Currently pursuing)

- **Sentiment Analysis -** I was responsible for model creation and validation, which was part of Relation Extraction project. During this period I worked on Different Classifiers, Clustering algorithms etc. Performed Sentiment Analysis to measure Brand performance (sentiment) and Topic Modeling to understand important topics.

- **Newsletters designing tool-** Designed a web application to draft-newsletters, which focused on XML-XSLT Transformation. An expert in writing XSLT style sheets and thus transforming the data, which would satisfy the vision statement of the project - 'Inform and Inspire'.

- **Technology Used:** Python along with Beautiful soup library for web data extraction. A simple and clean UI was completely developed using HTML 5, CSS and Java script.

Support Engineer Intern | Meltwater Software Pvt Ltd
Nov'15 – April 16

Project Name: New Meltwater Production Environment Monitoring

TASKS:

- Investigating on Kibana logs and resolving or reporting major anomalies.
- Configuring web-scrapping robots (tkRobot).
- Investigating GNIP (twitter) rules created at data pipeline. Configuring Ice-Rocket to fetch data from Facebook.
- Closely working with DevOps in patch release to production and managing testing team.
- Backfilling and monitoring Mongo DB at the time of cluster issues.
- Improve SLOs by automating the operational tasks using shell scripts and scheduling them using CRON jobs; Many instance of outrange failure of the application was mitigated.
- Styling of XML content using XSLT, HTML and JavaScript.
- Provide second and third level technical support to meet SLA's for all issues and their components.
- Server monitoring and maintaining using New Relic.
- Project planning and support using Jira Atlassian.
- Resolve Sev1/Sev2 incident tickets and service requests.

REFERENCES

Stanley Komban
Junior Data Scientist
Meltwater Software Pvt Ltd
stanley.komban@meltwater.com

Unni Chandroth
IT Director
Meltwater Software Pvt Ltd
unnikrishnan.chandroth@meltwater.com

Publications and Research Paper Implementations
July16 – Today

Paper Title : Surveillance System Based On Raspberry Pi
Journal: International Journal of Advanced Multidisciplinary Research (IJAMR), 2015
Link: **https://josevacha.github.io/online/** (Refer publications section)

- Recommender system using Low Rank Matrix Factorization.
- Anomaly Detection using (Multivarient and Normal Distribtion)
- Attended Fifth International Conference on Business Analtics and Inteligence.