

Aruna Arumugam

+61 452 443 130 | arumugama@student.unimelb.edu.au

Website: [aruna12.github.io/aruna-arumugam/](https://github.com/Aruna12) | LinkedIn: [linkedin.com/in/arunaarumugam/](https://www.linkedin.com/in/arunaarumugam/) | Github: <https://github.com/Aruna12>

SUMMARY

An aspiring data scientist currently pursuing Masters in Business Analytics to bridge the gap between technical and business skills to help organizations make data-driven business decisions. Experienced in text analytics and working with cross functional teams.

EDUCATION

THE UNIVERSITY OF MELBOURNE - MELBOURNE BUSINESS SCHOOL

Melbourne, Australia

Master of Business Analytics

Expected Dec 2020

- Statistical Learning, Optimization techniques, Causal Inference, Predictive Analytics, Text Analytics, Marketing Analytics

COLLEGE OF ENGINEERING, GUINDY

Chennai, India

Bachelor of Engineering in Computer Science and Engineering

Aug 2015 – May 2019

- Database Management, Data Structures and Algorithms, Machine Learning, NLP, Python programming
- Honours Grade: First Class Division

TECHNICAL SKILLS

- Programming languages:** Python, R, SQL, HTML, CSS, C, C++
- Database and ETL:** MySQL, Pentaho, Apache Spark, AWS
- Data Visualization:** Tableau, RShiny

PROFESSIONAL EXPERIENCE

PIPECANDY

Chennai, India

Data Science Intern

July 2019 – Dec 2019

- Scrapped news articles from various sources to create a dataset of 150k entries. Performed topic classification using SVM, Random forest and Naïve Bayes. Achieved 80% accuracy using RF. This automated the news article feature on the company website, eliminating the need for manual classification.
- Created a regular expression model to classify emails and urls from client's data sources consisting of 100,000 entries.
- Wrote script for data crawling using python's get request, soup and regular expressions to scrape reviews on various products of the client. Performed aspect based sentiment analysis on the reviews to provide insights about their customers.

Point105-AR

Chennai, India

Software Development Intern

May 2018 – June 2018

- Built a wordpress plugin as a prototype to display 3D images with glb and gltf extensions using wordpress API, php and three.js.
- Built an algorithm using python to optimize the decimation function in blender software.

PROJECTS

Identifying fake job postings ([project link](#))

June 2020 – July 2020

- Performed EDA and feature engineering on job description data set from emscad.samos.aegean.gr with 17k posts.
- Performed performance comparison between various feature combinations using logistic regression, random forest, naïve bayes, linear svc, KNN and concluded the tf*idf feature with logistic regression as the best model with 91% recall on fake job postings.

Finding trend of climate change ([project link](#))

June 2020

- Performed EDA and decomposition on time series data on Melbourne's monthly temperature data from July 1970 – February 2020.
- Performed a forecast for the next 10 years and concluded that there will be a small increase in the temperature.

Depression Detection on tweets using sentiment analysis ([project link](#))

July 2018 – Dec 2018

- Scrapped data from twitter using twitter API and tweepy based on hashtags. The model was trained on 5000 tweets.
- Cleaned and labelled data using textblob.
- Performed sequence encoding, decoding and classification/prediction using pytorch to say if there were signs of depression or not. Achieved a 92% accuracy with the model in detecting tweets containing signs of depression.

ADDITIONAL

- Achievements:** Winner of CODHER 2017 (Hackathon for girls conducted by ACM student chapter).
- Leadership roles:** Technical Head of Entrepreneurship Club (2017-2018), Vice Chairperson of ACM CEG.
- Certification:** Intro to Data Science in python, Applied ML with python, Applied Text Mining with python(coursera)