VIMAL DHARMALINGAM

Data Scientist

SUMMARY

- Total around ~4 years (3.8 years) of work experience as **Data Scientist** cum **Data Analyst** in Predictive Modelling - Banking, Industrial and Government sector and e-commerce and IOT.
- Rich work Experience in **Data Science with Machine Learning**, **Deep learning**, **Text Mining** and **NLP**.
- Extracting data from **S3 storage** and building predictive models in AWS Sage maker (Jupyterlab, Boto3) and reviewing the performance.
- Collecting the source data from S3 and processed trained and deployed in **AWS Sage maker** and uses deployed model for realtime prediction using **Serverless Lambda function** and **API Gateway**.
- Hands on experience of End to End Modelling, Data Mining, Data Analysis, Statistical Model, evaluating the models and data driven decisions.
- Experience on handling **structured, semi-Structured and un-structured data**.
- Hands- on experience in building predictive models using **Supervised and Unsupervised Machine learning algorithms**: Regression, Classification, Clustering and Natural language processing (Text Mining/Sentiment Analysis).
- Rich knowledge in Deep Learning methods such as Artificial Neural Networks, Natural Language Processing & Recommendation Systems.

WORK EXPERIENCE

Aug-2019 - Present	**Data Scientist, GovTech, Singapore.**
	Consultant, nsearch global pte. ltd.
Domain	IOT
Programming	**Python**, **scipy**, **numpy**, **spacy**, **SQL**
Tools and Products	AWS, Custom ML Model, **Sagemaker**, Athena, Personalise, **DynamoDB**, **S3**, Blazingtext, XGBoost, **Machine learning**, **Text Analysis**, NLP, JIRA, WIKI / CONFLUENCE, git / github, AWS codecommit.
Roles and Responsibilities	

- Responsible and Owned for End to End Machine learning use cases and Solutions using **Sagemaker**.
- Automated the **AWS** inbuilt **Blazing Text Neural network** model for text classification.
- **Automation**/Deployment of Machine learning models using **Sagemaker**.
- Realtime Inference **prediction** using **Lambda** function and **API Gateway**.
- Implemented training and **production** ready model using AWS machine learning services.
- Built and automated the Multilabel Classification problem to predict the categories of sensor using text input.
- Incorporating **ML/NLP** logics to improving the performance of elastic (**ELK stack**) Search Engine
- Automated the Text similarity model to find the application of sensors using **BERT** similarity model with semantic meaning
- Implemented the **Data mining** and **Data mapping processes** using NLP techs.
- Data standardisation of sensor devices field/ technical parameters using cosine similarity algorithm.
- Closely worked with the Data Engineer/backend and recommended which field to be in what **data format**. Also helped him to extract additional fields necessary for implementing **supervised** machine learning model.
- Was **SPOC** in the team and coordinated many modules/tasks and was helping delivery manager in achieving the same.

- Singapore
- +65 97155684
- vimald8959@gmail.com

CONTACTS

Address:

Jurong West ,

Singapore-640915.

SKILLS

Primary Skills and Work Experience

Machine Learning,Python,Data Science,NLP

Products and Tools with Work experience

AWS, Athena,Hadoop,Lambda,Teradata, Data Science, R, HBase, Hive, Hadoop, Spark-Pyspark,Teradata SQL

Secondary Skills and touch based in Work

REFERENCES

1)

Sathish Kumar

Data Science Manager at Hong Leong bank Berhad, Malaysia

Contact Details:As per Request

2)

Siah Shufen - Project Manager

Project manager at Govtech Singapore

Contact Details:As per Request

PERSONAL DETAILS

- Passport Valid till : 2024
- Passport No : M0766334
- DOB : Aug/1993
- Native : India, Tamil Nadu
- Current Location : Singapore
- Language : English, Tamil, Hindi

Projects/modules	• **Multilabel Classification** -To find the sensortype(Auto categorization)
	• **Short text Similarity** with Semantic meaning using Bert Similarity model
	• **Text matching** using Cosine similarity and word2vec.
	• Automation of **Data mining and Data mapping**
	○ **Spam Detection** to find the sensor article from the news .
	○ Realtime inference prediction using lambda endpoint

June-2018 - July-2019	**Data Scientist, Hong Leong Bank Berhad, Malaysia**
	consultant, Mission consultancy services

Domain	Banking (Marketing) Domain
Programming	**Python**, **Teradata SQL**, **scipy**, **numpy**, **pandas, R**
Tools and Products	**Tableau**, **SAS Base, IBM Watson , Google Analytics**
Roles and Responsibilities	• Responsible for End to End Modelling process and implement large scale solutions using **Python**, **jupyter**, **Teradata SQL** & **Base SAS**.
	• Predict the Customer service calls using ARIMA & SARIMA models
	• Built a model to identify Customers without PL has high propensity to buy PL (CASA & CC Base).
	• Built a model to **recommend merchants** by understanding customers implicit & explicit interests by **analyzing** their historical behavior.
	• **Sentiment Analysis**-Build a model to evaluate customer's feedback from multiple sources (Internet Banking, Social Media (Facebook, Twitter, Chat bot) & Call Centre.
	• Model is back tested by real time manual campaigns and compare the results of both (Response Rate of Manual Leads vs. Model Leads).
	• Experience in Personal Loan, Mortgage Loan, CC, CASA, INSURANCE etc.
	• Hands on experience on Attributes Selection for ML & Evaluating the Models.
	• Experience on Chatbot **(IBM Watson)** and **Text Analysis /NLP**.
Project 1:	• **Problem Statement:** To find high propensity customers who will buy Personal Loan in near feature for Marketing Campaigns
	• **ML-Tech**: Ensemble Models
	• **Solution: ensemble** model have been giving the **4% better lift** than manual campaigns.
Project 2:	• **Problem Statement:** To find high propensity to cross sell CC to CASA Customers
	to increase customer base and customer experience
	• **ML-Model:** Ensemble Models
	Solution: Neural Network model have been giving the **4.7% better lift** than manual campaigns.
Project 3:	• **Problem Statement:** Recommend the most appropriate merchant based on customer behavior & spending pattern**.**
	• ML-Model: Text Mining, TF-IDF Vectorizer, N-grams, Cosine similarity, Item based Collaborative Filtering
	• **Solution:** Using Item based collaborative Filtering Recommended the Merchants based on customer spending behavior
Project 4:	• **Problem Statement:** Developed python Machine Learning model to evaluated Bank's Chatbot (IBM Watson) responses
	• **ML-Tech:** TF-IDF, cosine similarity
	• **Solution:** Helps bank to retrain the bot on products where it was not answering accurately.

Nov2016 - Apr 2018	**Data Analyst/Modelling**
	Consolidated Construction Consortium Ltd

Domain	Industry
Programming	MS Excel, R.
Tools and Products	MS Excel, Powerpoint, MS Word
Roles and Responsibilities	Consolidated Construction Consortium Ltd, Chennai/Hyd (Reliance)

- Data extraction from Accident investigation reports and occupational reports.
- Closely deals with Safety HoD to optimize input features according to use case.
- Predicting accident occurrence and First Aid occurrence at site in advance.
- Mining the data for statistical Modelling using excel, Preparing daily safety analysis report, Finding possible ways to improve effective input features.

EDUCATION

Oct-2015 - July-2016	**Post Graduate Diploma in Management**
	Visvesvaraya Tedchnological University
	Graduated with 8.0 CGPA
Aug-2011 - Apr-2015	**Bachelors of Engineering**
	Anna University
	Graduated with 8.1 CGPA
2009 - 2011	⑂ **2011- Higher Secondary, state Board**
	Distinction in Physics, Maths.
	⑂ **2009- Higher Secondary, state Board**
	Distinction in Physics, Maths.

CERTIFICATIONS

- AWS certified Amazon Cloud Practitioner
- Pursuing AWS Machine Learning Speciality
- Stanford Machine Learning course by Andrew Ng [Coursera].
- Machine learning A---ZTM: Hands---On Python [Udemy].
- The Complete Machine Learning Course with Python [Udemy].
- NLP-Natural Language Processing with Python [Udemy].
- Elastic Search and Elastic Stack in depth and Hands On [Udemy].

PUBLIC CONTRIBUTION:

- https://pypi.org/project/py-custom-spellrectify/
- https://medium.com/@vimald8959/spam-detection-using-sagemaker-blazingtext-algorithm-b7e2ba2207c9
- https://medium.com/@vimald8959/sentence-categorisation-short-text-similarity-61bb88fae15e