

Rishin Rahim

Data Science and Machine Learning

📍 Bengaluru, India ✉ rishin07@gmail.com ☎ +91 9495952917 🔄 spacedrepetition

Education

Master of Science in Information Technology, June 2014

Indian Institute of Information Technology and Management Kerala

Bachelor of Technology in Information Technology, April 2011

College of Engineering Perumon | Cochin University of Science and Technology (CUSAT)

Work Experience

Tata Consultancy Services

5 Years , Oct 2014 - Present

Enview Research & Development - Research Intern

6 Months, Jan 2014 to May 2014

Projects

- **Intelligent Robotic Process Automation, TCS Robotics Bengaluru**
 - Built intelligent bots to extract data through various connectors, download various reports, search invoices and process the invoice to extract specific information.
- **Contract Digitisation - AI Platform for Legal Documents, TCS Robotics Kochi**
 - Built Entity recognition using NLTK library built-in methods and entity and classification model based on LSTM algorithm implemented using Keras Neural network, to extract metadata from Contract documents
 - Built classification models with RandomForestTree and other ensemble algorithms to recognise standard agreement clauses and values attached to it.
 - Built multilayered CNN models to identify handwritten clauses and attributes.
 - Training was performed as a batch process using apache airflow. Model was compiled and stored as pickle file. Prediction was done through a RESTful API built with flask.
 - Developed Dashboard /application interface including an ETL distributed queuing system, to monitor the different legal agreements, providing global overview of data through unique data visualisation.
- **UNSPSC classifier and Information Extractor, TCS Robotics Kochi**
 - Built a hierarchical product classifier based on RandomForestTree that classify products based primarily on their UNSPSC Code, brands and titles into a large taxonomy.
 - Built a Webscrapper using beautiful soup python library to extract data from relevant web URLs.
- **A Time-series Model for predicting CPU Utilization, TCS Assurance AI Lab Kochi**
 - Built the time series forecasting model using Auto Regressive Integrated Moving Average Algorithm (ARIMA).
 - The model predicted future change trend rules of the server CPU utilization which helped to improve the server performance and resource utilisation.
- **Root Cause Classification and clustering model, TCS Assurance AI Lab Kochi**
 - Built a Root Cause Analysis NLP classification model based on maximum entropy algorithm implemented using Apache OpenNLP.
 - Built clustering model based on lingo algorithm, implemented using Carrot², an Open Source Search Results Clustering Engine.
 - Developed topic modelling models using the Latent Dirichlet Allocation (LDA) from Gensim package.

- **Test Suite Optimization, TCS Assurance AI Lab Kochi**
 - Built a Test suite Optimisation NeuralNet Model implemented with a Keras neural network with three input layers used Adam optimiser, binaryCrossEntropy as loss function, RELU and softmax as activation functions.
 - Eliminated test case redundancy, improved test coverage and optimised the test effort.
- **Software defect prediction, TCS Assurance AI Lab Kochi**
 - Developed and deployed a Defect Prediction model that predicted the number of defects in future application release using multivariate subset linear regression and Correlation based feature selection.
 - Training was performed as a one off process and model was compiled and stored as pickle file. Prediction was done through a RESTful API built with flask.
- **Test case Similarity check, TCS Assurance AI Lab Kochi**
 - Built a similarity check model that measures the degree of similarity between different test step execution using cosine similarity based NLP algorithm.
- **Master's Thesis** :Threshold logic Object Detection using FPGAs .Thesis Advisor : Dr Alex P James
 - Two novel techniques for object detection are presented, one based on Resistive threshold logic and other based on the binary XNOR operation.The design and verification is done using Verilog hardware description language. Simulation and timing analysis is presented.The design is then synthesised and mapped into FPGA.
- **Neurosurgical eLog** : Advisors: Dr K.Srinivasan, Dr. Girish Menon
 - Web application to record and evaluate the surgical logs. Tolls used include web.py, gunicorn. Currently deployed and active in Sree Chithra Thirunal Institute of Medical Science & Technology.

Courses & Certifications

- **Machine Learning** Andrew NG, Coursera
- **Tensorflow in Practice Specialization:** deeplearning.ai, Coursera
Introduction to tensorflow, Convolutional Neural Networks in Tensorflow, Natural language processing in Tensorflow, Sequences, Time series and Prediction
- **Statistics with Python Specialization** - University of Michigan, Coursera
Understanding and visualising data, Inferential Statistical Analysis,Fitting statistical models to data

Skills

- **Machine Learning:** Classification, Clustering, Regression, Natural Language Processing, Recommender System, Ensemble algorithms, Dimension Reduction Algorithms, Neural Networks and Deep learning,Convolutional neural networks, sequence models
- **Statistics** : Probability, Descriptive statistics, inferential statistics, t-test, hypothesis testing.
- **Programming languages:** Python (Tensorflow, Keras, SciKit-Learn, Pandas, NumPy), R (ggplot2, Rcharts, dplyr), Java (Spring, hibernate, OpenNLP)
- **Application Development** : Python-flask, RabbitMQ, Celery, html5, CSS
- **Database** : Postgres, mysql
- **Visualization:** Python (Matplotlib, Seaborn, plotly), D3.js
- **ETL** : Apache Kafka