

ARVIND SINGH

arvind8198@gmail.com/(+91)8004440940

PROFILE SUMMARY

- More than three years of experience in Python, shell script, R, machine learning, Statistics, visualization, time series analysis
- The coursework relevant to my current pursuits include Python, shell script, NLP, image processing, Docker container,AWS services

SKILL SET

PROGRAMMING	Python,AWS Services-AWS Sagemaker, AWS Lambda, Amazon Opensearch, AWS Dynamodb,S3, SQS,SNS, Docker Container, R, Shell Script,c,c++, MS-Office, Ferret, CDO
COURSEWORK	Data Structures and Algorithms, Linear Algebra,Probability and Statistics, Data mining and AI techniques, Regression Analysis, Time series analysis, NLP, Theory of computation, & Numerical Analysis

EDUCATION

MAY '18	Dual Degree (BS - MS), Indian Institute of Technology, Kanpur <i>Major: Mathematics and Scientific Computing</i>	6.53/10* 5.2/10[†] (*-PG,†-UG)
JUNE '10	Saraswati Vidya Mandir Inter College, UK Board, Dakpathar, Dehradun	62.6%
JUNE '08	Asha Ram Vedic Inter College, UK Board, Vikasnagar, Dehradun	67.8%

WORK EXPERIENCE

Data Science Engineer, KareerMatrix, Noida *Oct'21-ongoing*

- working on a content based recommendation system for variety of content (video, audio, images, books)
- worked on documentation, design, development, deployment and testing of the **recommendation system**
- Developed an **NLU** (Natural language understanding) based events search engine using **BERT** and **Amazon Opensearch**
- Developed an **NLP** topic model using **LDA** to extract keywords from books to create word embeddings for books

Project-Scientist-B, INCOIS, Hyderabad

Mar'19-Oct'21

- Worked on spatiotemporal deep learning model for SST field prediction using time series satellite data
- Major responsibilities include Automation of data Quality control, format conversion,ocean model data analysis and validation
- Build a **VAR** (vector auto regression) model to forecast MLD and analyze contribution of DSWRF in forecast error variance of MLD

FREELANCING WORK

Work done as a freelancer (<https://github.com/dislu/freelancing>) *Jan'20-Present*

- Constructed classification tree models to predict student dropout and compared models on appropriate validation metrics
- Build a **NLP** model to map sentiments in student notes and then discover topics using **LDA** algorithm
- Did social network analysis of a school classroom and then analyzing both centrality measures and clusters within the network

PG PROJECT

A finite volume study of the blood flow in total cavopulmonary connection (TCPC) *Jan'17-Apr'17*

- Studied the flow reversal dynamics of fluid in the academic version of **TCPC** arrangement
- Developed many models by varying offsets of TCPC arrangements and simulated the model using **Ansys Fluent**

ACADEMIC PROJECTS

The Churn Game (course project, Summer Semester) *May'16-Jun'16*

- Developed predictive models for predicting customer churn in the Telecom Industry
- Determined the workability of the model and estimated the profit after implementation of the suggestions

WLAN Fingerprint Based Indoor Localization (course project, 6thSem)

Aug'15-Nov'15

- Build a Regression model to detect indoor location of a user (latitude, longitude) using RSSI and WAPS signals
- Achieved **94.3%** Adj R-sq and **2.4%** max error with validated data

Grouping and Reporting loan applicants data (course project, 6thSem)

Jan'15-Apr15

- Used data mining techniques like **PCA** and **K-mean clustering** for grouping and Reporting loan applicants data
- Efficiently separated out loan applicants into distinct categories and various plots were produced to visualize the data

ADDITIONAL INFORMATION

Extracurricular Activities *Mar'13*

- Secured a place in final 5in a technical event ECDC(electronics circuit design competition)Of electronics club
- Completed project using only 2ICEs and was cheapest gaming console in the competition

Hobbies include Playing cricket and Watching movies