Nishant Arora

CONTACT

nishant.arora.682@gmail.com
www.nishantarora.ml
+91-9929061890

EDUCATION

B.E. (Hons.) Mechanical Engineering and M.Sc. (Hons.) Mathematics
2014 — 2020

Birla Institute Of Technology and Science,Pilani

CGPA : 5.47

Class XII (A.I.S.S.C.E)
2013 — 2014

GMSSS, MHC (Chandigarh)

Percentage: 88%

Class X (A.I.S.C.E)
2011 — 2012

Bhavan Vidyalaya, Panchkula

CGPA: 10

RELEVANT COURSEWORK

Computer Programming, Data Structures and Algorithms, Object Oriented Programming, Machine Learning, Data Mining, Information Retrieval, Deep Learning Specialization, Number Theory.

WORK EXPERIENCE

Ecom Express
Jan 2020 — Present

Data Science Intern

Working with both Data Engineering and Data Science applications primarily using Apache Spark, Python, AWS

couture.ai
July 2019 — Dec 2019

Data Science Intern

Worked on Apache Spark and Scala (and occasionally with Python) to build scalable Data Science pipelines. Prediction,Topic Modeling, Recommendation, Feature Selection and Engineering, NLP were some of the pipelines I built. Also researched about AutoML and the implemention of such a platform.

nearbuy.com
May 2018 — July 2018

Software Development Intern

A/B Testing Framework based on Bayesian Statistics with a user friendly UI that shows test results in real time. Also devised a new Listing Ranking algorithm using Machine Learning which optimizes various objectives.

nearbuy.com
May 2017 — July 2017

Software Development Intern

Natural Language Processing Project : Designed a free text search parser using Apache Spark in Java. The classification model used was a Random Forest Classifier and the accuracy reached was 94%.

DoSelect   Aug 2016 — Feb 2018

Problem Setter

Set Up algorithmic problems for programming contests at DoSelect. Conducted 5 programming contests and hackathons for DoSelect.

H.M.T. Business Group   May 2016 — July 2016

Summer Intern

Devised and Implemented an algorithm for the automatic generation of unique supplier code using C++ and MySQL which marked the end of the manual system which was in existence from years.

PROJECT WORK

- Worked on the implementation of an automated scalable end to end forecasting pipeline from fetching the raw data everyday to delivering the forecasts, using Apache Spark, Arima, Python, S3, MySQL.
- Implemented a NLP Pipeline in Apache Spark and Scala which does various basic text cleanups to spell correction, part-of-speech based filtering, phrase extraction.
- Worked on the implementation of a pipeline which does automated feature selection from data.
- Implemented sales prediction pipelines using xgBoost and lightGBM models in Apache Spark and Scala.
- Worked on the implementation of recommendation and similar products pipeline using Topic Modeling in Apache Spark.
- Implemented a visual search engine within elasticsearch which uses image vector extraction using the Resnet-18 Model, self implemented vector to string encoders and ranking algorithms.
- Worked on implementing pipelines which ingest raw as well as processed data from MySQL database to the S3 datalake which use various connectors between Apache Spark code, S3, MySQL and Python.
- Analysing Data by fitting various probability distributions to it using PySpark and Scipy.
- Locality extraction on a Pincode level by clustering on latitude and longitude values using Apache Spark and Scala.
- Implemented various 2D packing algorithms in MATLAB and Python.
- Scraped advanced football statistics from understat.com to a Pandas dataframe and used that to cluster the players based on their goal scoring and playmaking capabilities.
- Developed a Flask Webapp which analyzes the person's intelligence levels through the input fingerprint ridge count and input values and generates a report of the same.
- Worked on the Enron Dataset and tried to identify various POIs using various Machine Learning techniques.
- Analysed the Football World Cup 2018 Data using various Machine Learning algorithms.
- Made an Object Following Robot in first year using MATLAB and Arduino.

ACHIEVEMENTS AND EXTRA CURRICULAR ACTIVITIES

ACM ICPC Asia Amritapuri Site Regional Contest 2016 Participant
Team Name: treeBitsians

Online Judges:

- Hackerrank Algorithms Domain World Rank: 200, Percentile : 99%
- Codeforces Rating: 1614, Problems Solved > 500
- SPOJ World Rank: 2500

Won the award for the Best Intern at nearbuy.com in the summer of 2018.

Secured 3rd position in CodeGenesis 1.0 on Hackerearth.
Secured 4th position in College Hackathon on Hackerearth.

Secured 8th position in Inter University Programming Contest 2016 at Plinth TechFest.
Microsoft codefundo++ 2018 participant.

Gold Medals in various District and State Level Chess Competitions and Participation in North-Zone and Nationals.

Interested in Competitive Coding, Data Science and Artificial Intelligence. Hobbies include playing Chess and Football, Puzzle Solving, Reading, Watching Movies and Gaming.

POSITIONS OF RESPONSIBILITY

- Core Member at BITS-ACM (Association of Computing Machinery)
- Event Coordinator of Programming Contests conducted by BITS-ACM
- Core Member of the Mathematics Association
- An Active Students Union Volunteer
- Managed LAN Sharing Facility for the Whole Campus

TECHNICAL SKILLS

Basic Skillset : Apache Spark, Scala, C, C++, Java, Python, GoLang, MATLAB, R, Spark Streaming, mmlspark, Smile scala, MongoDB, AWS, Tensorflow, Keras, PyTorch, Elasticsearch, Machine Learning, Flask, Heroku, fpdf, Pandas, BeautifulSoup, Number Theory, SQL, PL/SQL, Git, Apache Kafka, A/B Testing, HTML/CSS, Javascript.
Operating Systems: Linux, Mac OS, Windows
Software: Intellij, MySQL-Workbench, IBM SPSS,Visual Studio, PowerBI, STS, Metabase, Autocad, Solidworks, DartFish

ONLINE CERTIFICATIONS

- Introduction to Machine Learning, Udacity
- Neural Networks and Deep Learning, Coursera
- Algorithms : Design and Analysis, Stanford University
- Improving Deep Neural Networks, Coursera
- Algorithms : Design and Analysis II, Stanford University
- SEO for Beginners, Udemy
- Structuring Machine Learning Projects, Coursera
- Convolutional Neural Networks, Coursera
- Sequence Models, Coursera
- Text Retrieval and Search Engines, Coursera