nagarro
THINKING BREAKTHROUGHS

Pranav Malik

Associate, Technology- Big Data

Experience summary

Pranav has 25 months of experience in the software industry. He has experience in providing Big Data solutions using technologies such as Java ,Spark, Python, Scala, etc. He has skills in working on various domains such as Big Data, analytics, Machine Learning, Web Development via NodeJS and Angular.

Contact Details

Email – pranav2134malik@gmail.com

Phone no- 999318323

Trainings and certifications

- Java and Data structures, Python and Data Structures Coding Ninjas
- Apache Spark 3 with Scala, Udemy
- Apache Spark Streaming ,Big Data Udemy
- Taming Big Data with Spark Streaming and Scala, Udemy
- Data Analyst Nanodegree, Udacity(Cleared Level 1,2,3)
- Node JS,Angular,Vue.Js Professional course, Udemy
- Taming Big Data with Spark Streaming ,Udemy
- Scala for Big Data ,Udemy

Skill set

Programming and scripting languages	Python, Scala, R, Java, Pixi.js, Papyrus, JavaScript, Node.js, Angular, React Native, UnrealScript, C++, TypeScript, OpenGL, Dailogflow, Vue.js, Vanilla.js, Sqoop, Power BI, SPSS
Frameworks, tools, and libraries	Splunk, Hadoop, Spark, Storm, Cassandra, RapidMiner, Unity 3D, Unreal Game Engine, Phasor, Neural Networks, OpenCV, TensorFlow, NumPy, Keras, Panda, Hive, MapReduce, Amazon AWS-EC2,Lambada,Athena,Hadoop, PySpark.
Servers and platforms	REST Services, JSON, XML, e-commerce server backend using SQL Server,StreamSets,Kubernetes-Docker .
Databases and BI	SQLite, MongoDB,Cassandra,MongoDB, Neo4j Graph Data Bases
Devices and OS	Windows OS, Oculus Rift, IoT, Arduino, Raspberry Pi, Linuz

Education

- Bachelor of Technology, Computer Science, University School of Information and Communication Technology, Guru Gobind Singh Indraprastha University, Delhi, India

Recent projects

Price Master Product(Indus Valley Partners-Noida)

Client	A company with a client base consisting of the most successful hedge funds and private equity firms in the world.
Project overview	Project involved developing an OCR using deep learning algorithms and natural language processing to analyze client invoices and adding a submodule in price-master product using Django-Python framework. It involved: – Optimizing the existing search pattern token files of price master using state of the art solution in React and made suitable API for its deployment – Developing IVP transport application, providing additional functionalities for the existing transport application, and receiving appreciation from the Manager for this project – The project involved developing recommendation system with long term and short-term profile on Hadoop and OLAP DB for recommendation system. It involved forming multi node Hadoop cluster for handling and storing huge amount of data for data analytics. The task required integrating Sqoop to transfer terabytes of data from Vertica to maintain historical data delivering 300 odd reports in worlds different calendar in a Critical Dataware housing project. The project also required the following: – Performance tuning of the jobs (ETL) and DB Scripts with many scenarios such as manual projection design, query profiling, DBD etc. – Developing complex framework which handled inter-related query, waiting for each other processes completion (interdependent query mechanism – Integrated Sqoop to transfer data from source – Developed complex merge process to merge output from different system – Profiling and creating mapping document – Developing Hive script to generate internal and external table at runtime
Technologies	Python, Dailogflow, Neural Networks, Django, MongoDB, Angular 6, Java, Scala, R, SPSS, Data Miner, Git, Graph Databases,Cassandra

Responsibilities	• Involved in analysis of required functionality, task estimation, and allocation to the team • Designed application architecture and database • Worked on implementation of key modules and UI creation

KGH Customs(Nagarro)

Client	**Through a wide range of services, KGH optimizes your trade and customs management – helping improve trade performance, operational efficiency, risk management and customs duty control.**
Project overview	Project involved developing an Online Web portal for Custom documents uploading and processing via the government ,broker, . It involved: – Addition of a Custom Good Item description module , provided a dynamic loading platform via Vue, NodeJS – XML based document type database was used to handle and query over large chunks amount of Data. – Deployment was done via EC2,Lambada – Fixing frontend bugs via vue.js, html,css
Technologies	NodeJS, Amazon AWS- EC2,Athena,Lambada,VueJS,C#
Responsibilities	• Involved in analysis of required functionality, task estimation, and allocation to the team • Designed application architecture and database • Worked on implementation of key modules and UI creation

Smart India I4C-NCL Pune (Deloitte)

Client	***NCL PUNE is a Research and Development and Consulting organization with a focus on chemistry and Chemical Engineering . It was an Joint Venture of Deloitte and SIH i4c-NCL Pune***
Project overview	Project- Risk assessment of fire hazards in coal mines using data analysis and machine learning – A watch was made as a final project with the help of National Chemical Laboratory Pune(NCL-Pune) which is fitted with the technology to detect early fires before its occurrence A special care has been taken with respect to the issue of false alarms – A recommendation system written in Cassandra and Java, with front-end being in Angular 8, with basic functionality of elastic search

	– Exposed Machine Learning Model via a Flask Application
Technologies	Python ,IOT,Adruino ,Django Framework, Cassandra,Java,Angular 8,Flask
Responsibilities	• Involved in analysis of required functionality, task estimation, and allocation to the team • Designed application architecture and database • Worked on implementation of key modules and UI creation • An award of INR 1,00,000 was awarded.

Big Data Practice(Nagarro)

Client	*Involved in Various Sub-Projects(POC)*
Project overview	**Zain Networks** – Anomaly Detection in Telecommunication Systems , using Markov Chain+Ann model , Working in Geospatial Data using OpenLayer,Hadoop and D3.js, Neo4j Graph Databases) – Elastic Search feature was used to ingest data from producers connected via Kafka Message Queue. – Hadoop was used in Data Lake feature of the pipeline where a data generator was scheduled via Airflow to dump data at a Particular Time Interval. – Data Dumping onto open layer was done by PostgreSQL and Python Script – Job Scheduling was done via Airflow – Optimized Anomaly Detection Process by adding Forecasting power using Supervised Learning using LSTM-Autoencoders – Modified the WMS Open layer to WFS which enable querying, non-static data loading, Auto-Refresh Feature, Dynamic Data Loading – Data Modeling was done in D3.js. – Visualization was done Via Kibana Charts using Dynamin Data uploaded by Hadoop file system. – Data Pipelines were made via Stream sets. – Used Sharding-Load Balancing Concept for better utilization of Clusters for memory and resource allocation. **Amway Data lake** - Data lake built on top of AWS infrastructure using S3, Kinesis for data ingestion and consuming purposes, Data churning and Business logic implemented by Go Lang and Python Lambda Functions - Live Order handling Mechanism – Handling 10^7 orders everyday - Recommendation System made by using Collaborative Filtering Using Boltzmann Machine – Providing Recommendation to over 1 million Users. - Search Raw Orders facility for reconciliation purposes
Technologies	PySpark,StreamSets,Kafka,Cassandra,Hadoop , Machine Learning ,ELK stack, Airflow
Responsibilities	• Involved in analysis of required functionality, task estimation, and allocation to the team

	• Designed application architecture and database • Worked on implementation of key modules and UI creation • Individually Handled the Machine Learning Part and Hadoop Cluster Set in StreamSets