**Raship kumar**

**Senior Data Scientist**

**(Service delivery manager)**

**[+91 8285088151 ]&[**[raship.kumar@cognizant.com](mailto:raship.kumar@cognizant.com)]

[https://www.linkedin.com/in/raship-kumar-9292a157](https://www.linkedin.com/in/raship-kumar-9292a157/)

**8 years of Expertise in developing & deploying Intelligent, automated business decision making applications using data science & analytical skills, resulting in business process and value optimization.**

**SKILLS**

|  |  |
| --- | --- |
| Operating Systems | Linux/UNIX, Windows 2000/XP/NT |
| Programming Languages | R, Python, SQL, Perl, UNIX Shell Scripting |
| Concepts | Statistics, EDA, Machine Learning, Predictive Analysis, Time Series analysis, Cluster Analysis, GARCH, Deep learning, Survival analysis, Feature engineering & selection, Tableau, NLP |
| Tools & Packages | Spider, SVN, WEKA,Nginx, flask, Waitress, Lime, KNIME, MS Office (Word, Excel & PowerPoint |
| Certification | Post Graduate Executive Certificate Program in Data Science(IMT Ghaziabad) |
| ML model deployment platforms | HEROKU, GCP, AWS |

**Data Science projects**

**Product recommendation system(Sysco)**

Cognizant Technology Solutions Sep’22 – May’23

* Designed and developed an AI solution to identify most relevant substitute product incase ordered product is not available in inventory
* Recommendation system helped client to reduce cancelled orders thus saved him lot of cost.
* Relevancy was based on similarity matrix using cosine similarity metrics.
* Implemented various NLP techniques in python
* Used flask framework for exposing model results.

**Email Management System(Walmart)**

Cognizant Technology Solutions Jan’22 – Jul’22

* Lead the development of a web app which parse/process outlook mails and then classify, priorities, generate and send auto responses by selecting mail routes as per the business logic.
* Developed architecture for the application
* Implemented text classification model for classifying mails into various categories.
* Implemented NER model for fetching entities of interests.
* Developed backend using Python-Flask web services framework.

**Employee Attrition prediction model**

Cognizant Technology Solutions Sep’21 – jan’22

* Performed EDA to understand data distributions and applied various statistical measures for feature engineering and feature selection.
* Used various methods for Hyper parameter tuning and optimization
* Used Xgboost algorithm for developing final binary classification model, based on EDA results.
* Build an autonomous ML system using Python and flask framework which outputs easily understandable business insights supported by various visual artifacts and statistics using raw data.
* Done Survival Analysis on HR data for estimating Survival curve/function, calculating employee survival probabilities as per their employment duration.
* Compared survival duration of different employee age groups to know which age group have more tendency to attrite.
* Prepared Interactive data analysis dashboards using Python library Streamlit and Panel
* Presented modeling and analysis results to various stakeholders.

**Claims Forecasting model**

Cognizant Technology Solutions Jun’21 – Aug’21

* Used Sarimax models for modeling claim settlement time series data using Python.
* Observed components like Trend, seasonality, cycles & white noise for optimizing predictions.
* Utilized forecasting results for predictive maintenance by analyzing future medical claim settlement flow among different claims categories.

**Credit risk modelling**

NIIT Technologies(Santender) Sep’19 – Jan’20

* Designed and implemented a machine learning system that predicted Expected loss with complying Basel and IFRS9 standards..
* Done effective data processing and feature engineering in order to interpret PD model and use results for score card generation.
* Developed Machine Learning model using Logistic regression used for application and behavioral scorecard development.
* Helped client in increasing revenue generation by minimizing Total Expected Loss (EL) using historical data including macro economic variables for developing for ttc-PD and pit-PD(Probability of default) models.
* Performed Roll rate analysis for deciding definition of defaults.
* Applied PD model validation for ensuring model meets specified requirements
* Back tested rating system by evaluating characteristics like Stability, Discrimination and Calibration.
* Monitored model’s input and output using various statistical measures on regular bases to monitor model’s performance over time.
* Develop segment and account level CCAR stress loss models.
* Applied Garch for predicting the volatility of returns on financial time series based return data.
* Developed a quantitative framework for investment portfolios that optimizes relation between risk & returns.
* Performed Mean-variance analysis for making better investment strategies.
* Presented modeling and analysis results to various stakeholders.
* Econometric modeling-driven stress loss process (data collection, data integrity pre-processing, segmentation, variable transformation, variable selection, econometric model estimation, sensitivity testing, back testing, out-of-time testing, model documentation, and model production implementation) Collected, combined customer demographic, financial well-being and linkert scale baseddata and prepared, cleaned, transformed to make data fit for modeling.
* Prepared multi-classification model using ANN algorithm for categorizing customer’s into appropriate risk categories.
* Utilized segmentation/clustering algorithms for performing customer profiling offering relevant offers/promotions to customers.

**Customer’s Risk Profiling**

NIIT Technologies(Santander)Mar’20 – Oct’20

* Collected, combined customer demographic, financial well-being and linkert scale baseddata and prepared, cleaned, transformed to make data fit for modeling.
* Prepared multi-classification model using ANN algorithm for categorizing customer’s into appropriate risk categories.
* Utilized segmentation/clustering algorithms for performing customer profiling offering relevant offers/promotions to customers.

**Return to Origin (RTO) Prediction Model**

Commdel Consulting Services Pvt. Ltd.(Homeshop18)Oct’18 - Jul ’19

* This binary classification was aimed to identify online orders having high probabilities of return
* Conducted EDA using various statistical measures to study relevant features and their interactions contributing to high probability return orders.
* Applied feature engineering & feature selection utilizing statistical measures to produce best model.
* Prepared Binary classification model for predicting orders with high probability of return. So that vendor can take simple proactive steps to stop order being returned.
* Used clustering algorithms for offering relevant offers/promotions to correct set of customers.
* Used lift curves to compare results with and without predictive model.
* Engineered features RFM features for customer segmentation/clustering.
* After using model return orders were dropped to ~11%.

**Product Substitition engine**

Tech Mahindra(Sysco) Oct’17 - Apr’18

* Created clusters of products using product data.
* Developed product similarity matrix based on cosine distance between product features in resulting clusters.
* Applied business rules on recommended products for adhering to recommendation policy.
* Deployed model as Flask web service in Python.

**Anomaly detection to track fraudulent patterns**

Tech Mahindra(AT&T) May’16 - Feb’17

* Analyzed unstructured & structured data and created the Anomaly detection tool using Isolation forest.
* Model can analyze financial data related to any process such as claims, payments, reserves, transactions etc.

**Customer Churn model**

Tech Mahindra(AT&T) Sep’15 - mar ’16

* Developed Machine leaning binary classification model for predicting customers having high probability of churn.
* By retaining probable churned customers client management was able to maintain steady customer base. Thus model helped in increasing life time value.
* Utilized Clustering algorithm DBScan for clustering customer data
* Developed and tuned regression model for predicting Customer lifetime value with least possible RMSE.
* Treated data for outliers/anomalies to point out suspicious data points and understanding data distribution better for accurate predictions.
* Combined solution helped management optimize revenue by 7%.

**EXPERIENCE**

Cognizant Technology Solutions as Analytics Delivery Manager May’21- present

NIIT Technologies as Sr. Data Scientist Jul’19 – dec’20

Commdel Consulting Services Pvt. Ltd. as Sr. Data Scientist Oct’18 - Jul ’19

Tech Mahindra, Noida as Data Scientist Sep’15 - Jun’18

HCL Technologies, Noida as Tech lead Jan’15- Aug’15

India Bulls Technology, Gurugram as Data Scientist/Automation Engineer Oct’12-Dec’14

Headstrong Services India Pvt. Ltd., Noida as Senior Associate Feb’11-Sep’12

Towards Vision Technologies Pvt. Ltd., Gurugram as Software Engineer Sep’08-Oct’10

**EDUCATION**

Completed **B.Tech. in Computer Science** from, Kamla Nehru Institute of Technology, Sultanpur(Uttar Pradesh, India) in the year 2008.

**Personal Details**

* Date of Birth: 19th April 1985
* Languages Known: Hindi and English
* Address: Y1-1204(A), Eros Sampoornam, Sector-2, Greater Noida – 201306
* Hobbies: Playing table tennis, Reading technical blogs, listening music