

NISHIT CHAUDHRY

Data Scientist

+1-716-730-9191 • nishitch@buffalo.edu • <http://www.linkedin.com/in/nishitchaudhry/> • <https://nishitchaudhry.github.io/>

EDUCATION

| | |
|--|-----------|
| State University of New York at Buffalo – Master of Science, Management Information Systems (Analytics) (STEM) | Jun. 2022 |
| IIIT, Bangalore – Post Graduate Diploma, Data Science | Jul. 2021 |
| Shiv Nadar University, India – Bachelor of Technology, Electrical and Electronics Engineering | May 2019 |

WORK EXPERIENCE

| | |
|---|-----------------------|
| Cognizant Technology Solutions, Data Scientist – Pune, India | Jul. 2019 - May 2021 |
| <ul style="list-style-type: none">Tech Stack – Python, AWS, Scikit-Learn, TensorFlow, Spacy, NumPy, Pandas, Matplotlib, Agile-ScrumImplemented machine learning pipelines such as data preprocessing, exploratory data analysis, feature engineering, model building, model evaluation, and model deploymentCollaborated and developed optical character recognition as-a-service on hand-written and printed text in tabular format with word-level accuracy of ~85% by leveraging YOLO, OpenCV to increase efficiency of automated hand-written document readerPerformed different iterations of abstractive and extractive text summarization by utilizing Natural Language Processing transformers to attain ~87% accuracy in comparison with human-level text summarizationPresented and designed Rasa conversational chatbot for FAQ use-case and proposed advantages of including Rasa X for conversation-driven development to improve efficiency by ~1.5 timesContributed to enhancing Rasa chatbot into multi-lingual chatbot to converse and respond with users in 6 European languagesAchieved accuracy of ~82% in multi-class support ticket classification model using data augmentation techniques to automate ticket support generation and increase process efficiency | |
| Cognizant Technology Solutions, Data Science Intern – Pune, India | Jan. 2019 - Apr. 2019 |
| <ul style="list-style-type: none">Tech Stack – Python, Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, Excel, Jupyter Notebooks, Agile-ScrumPresented and prepared ARIMA model on sales and revenue time-series forecasting for European telecom giantObtained accuracy of ~91% with confidence interval of ~95% in predicting revenue | |

ACADEMIC PROJECT EXPERIENCE

| | |
|--|-----------------------|
| Buffalo Crime Rate Response Prediction (Regression, Python, Tableau, Tableau Prep Builder, SAS Eminer) | Oct. 2021 - Nov. 2021 |
| <ul style="list-style-type: none">Prepared, and visualized crime dataset with 48k records across 45 dimensions using Tableau Prep Builder and Tableau DesktopCollaborated and predicted crime incident response time of law enforcement officials in Buffalo, NY using machine learning models such as OLS, Random Forest, Decision Trees built on SAS E-miner and Python | |
| Brain MRI Style Transfer (GANs, Python, TensorFlow) | Jun. 2021 - Jul. 2021 |
| <ul style="list-style-type: none">Implemented GANs model to generate artificial MRI images of different contrast levels from existing brain MRI scans to enhance accuracy of diagnosis by providing practitioners with comprehensive understanding, saving cost and time by 50% | |
| Skin Cancer Detection (CNN, Python, TensorFlow) | Mar. 2021 - Apr. 2021 |
| <ul style="list-style-type: none">Detected Melanoma type skin cancer using Convolutional Neural Networks on dataset with 9 skin cancer class typesAchieved accuracy of ~73% after regularization and data augmentation for class imbalance | |
| Telecom Churn Case Study (Classification, Python, Scikit-learn) | Jan. 2021 - Feb. 2021 |
| <ul style="list-style-type: none">Developed predictive models to identify customers at high risk of churn and identify main indicators contributing to churnConstructed classifier models such as Logistic Regression, Random Forest, Decision Trees, SVM with f1 score of 0.93 | |
| House Pricing Prediction (Regression, Python, Scikit-learn) | Dec. 2020 - Jan. 2021 |
| <ul style="list-style-type: none">Analyzed data with 80 dimensions and built ridge regression model to project house pricesReported significant features contributing to house pricing and obtained r-squared score of ~81% | |
| Lead Scoring Case Study (Classification, Python, Statsmodels, Scikit-learn) | Nov. 2020 - Dec. 2020 |
| <ul style="list-style-type: none">Created logistic regression model with f1 score of 0.77 to assign lead scores between 0 - 100 to target potential customer leads | |

SKILLS AND CERTIFICATIONS

Knowledge Areas: EDA, Predictive Analytics, Data Visualization, Data Mining, Statistical Analysis, Machine Learning, Time Series Forecasting, Neural Networks, Hypothesis Testing, Data Scraping, Experimentation

Languages and Tools: Python (Pandas, Matplotlib, NumPy, Scikit-learn, TensorFlow, Spacy), SQL, Tableau, SAS Eminer, AWS, Google Analytics, Microsoft Excel, Hadoop, Jupyter Notebooks, Git, NodeXL

Certifications: AWS Certified Cloud Practitioner, IBM Data Science Professional Certificate, Machine Learning A-Z: Hands-On Python In Data Science, Hacker Rank SQL, UC Davis SQL for Data Science