EASHAN ARORA

Chicago, Illinois| earora3@uic.edu | (312) 391-1892
Linkedin | Portfolio

EDUCATION

University of Illinois Chicago December 2022
Master of Science in Management Information Systems GPA:3.78/4.0

- Relevant Coursework: Advance Database Management, Statistics for Management, Data Mining in Business, Machine Learning

Jaypee Institute of Information Technology, Noida, India June 2020
Bachelor of Technology in Computer Science

- Relevant Coursework: Software Engineering, Computing for Data Science, Database Systems

TECHNICAL SKILLS

- **Programming Languages**: Python, R, SQL
- **Machine Learning**: Regression, Decision tree, Random Forest, K-means clustering, NLP
- **DATABASES**: MySQL, SQL Workbench
- **OTHER TOOLS**: Statistical Tests (t-test, Regression Analysis), Tableau, Git, Jupyter

PROFESSIONAL EXPERIENCE

LUMIQ, NOIDA *(An Analytics startup in INDIA)*
DATA SCIENTIST July 2020-July2021

- Created an Interaction Engine for a Banking Firm in India using an open source Chatbot Platform 'RASA' to resolve customers day to day queries which came in the form of human utterances.
- Customized the initial RASA pipeline based on the client use case and assessed the use of 6 modules proposed in the conversational AI using NLP techniques.
- Engineered a Video/Image Analytics project to extract relevant information from 100+ legal documents using NLP techniques and face detection of customers using convolution neural nets.

RADIX INFO SOLUTIONS
Data Science Intern

- As a data science trainee coded on different data wrangling techniques using Python and learnt statistical test methods for data mining used for calculating significance of each attribute in a particular dataset.
- Prepared data analysis for a Zomato (a food delivery app) dataset retrieved from Kaggle.

ACADEMIC PROJECTS

Natural Language Understanding Pipeline

- Created an automated NLU pipeline which includes 5 modules aiming to provide a semantic way for user utterances.
- Intent classification (context of the sentence), slots, query type (type for the sentence), query correctness (Probability for the query being correct) and sentiment for any user utterance through NLP and Simple transformers framework.
- Overall Probability scores for models were greater than 95 percent.

Prediction for Employee attrition in the dawn of recession

- Performed exploratory Data Analysis on the IBM Dataset of around 2000 instances, applied Feature engineering, Feature encoding, Attribute selection.
- Machine Learning techniques including Data cleaning, Model Selection using decision tree, random forests.
- Achieved over 80 percent AUC-ROC scores for this kaggle project.

Data mining approach to predict Forest fires

- Designed a thorough comparative study to showcase the relative and absolute efficacy of 5 different Data Mining techniques in predicting the expanse of forest fires.
- Overall Decision tree performed the best with an accuracy of 98 percent.

Research Publications

- Eashan Arora, Sakshi Mishra, K.Vimal Kumar, Pawan Upadhyay, "Extending Bidirectional Language Model for Enhancing the Performance of Sentiment Analysis", published at Part of the Lecture Notes in Electrical Engineering book series (LNEE, volume 643).