NETRA PATHAK

+1 (732) 675-4799 • netra@uw.edu • http://www.linkedin.com/in/netra-pathak • https://github.com/netrapathak

EDUCATION

University of Washington Seattle, WA
Master of Science in Data Science, GPA: 3.7 *Sep 2018 – Jun 2020*
- Relevant Coursework: Business Intelligence and Data Warehousing, Data Science, Interactive Information Visualization.
- Graduate Assistant for 'Core Methods in Data Science', Winter 2019; 'Python for Information and Data Science', Autumn 2019.

Ahmedabad University Ahmedabad, IN
Bachelor of Technology in Computer Science *Jul 2013 – May 2017*
- Relevant Coursework: Data Analytics, Machine Learning, Software Engineering, Databases, User Centered Design.
- Leadership: Student Training and Placement Cell Coordinator; Member of Executive Council, IEEE student branch.

WORK EXPERIENCE

Human-Centered Data Science Lab, University of Washington Seattle, WA
Research Assistant *Sept 2019 – Present*
- Investigating online fanfiction communities using visual analytics to develop a tool to help authors reflect on their writing.
- Engineered interactive Tableau dashboards to highlight trends and summary statistics of weekly reviews received over a year, resulting in appreciation from team. Leading focus group interviews and usability testing to evaluate the tool.
- Enhanced the dashboards using NLP techniques: created word embeddings to visualize clusters on TensorFlow Projector and used tf-idf feature selection and term normalization, resulting in 15% improvement of the feedback tool.

VMware Palo Alto, CA
Data Science and Analytics Intern *Jun 2019 – Sep 2019*
- Analyzed 2000+ tickets in collaboration with cross-functional Incident Management teams to identify top 5 customer pain points in VMware Cloud (VMC) on AWS. Involved in full analytics lifecycle from idea to analysis, decision and execution.
- Devised a prototype for time series analysis of ticket resolution using Prophet to identify anomalies and detect trends in capacity usage of VMC, leading to 10% increase in efficiency of Engineering Escalation Management.
- Designed a dynamic webpage using HTML and JavaScript to help business team analyze customer interactions, utilization and remediation, saving 3-man hours per day.

Arya.ai Mumbai, IN
Data Scientist *Jul 2017 – Aug 2018*
Arya.ai is one of the leading AI startups in India, offering tools to build, manage and scale complex deep learning systems.
- Trained and deployed deep Convolutional Neural Networks for Automated Cheque Truncation System, devised for top banks in India to extract key details like account number and date details on cheque, reducing manual work by 32 cheques/second.
- Led stakeholder communications and automated the evaluation pipeline by developing API that generated verification metrics.

Sty-Labs Info Solutions Pvt. Ltd. Mumbai, IN
Computer Vision Intern *Jan 2017 – Apr 2017*
- Modeled CNN in Keras with 80% accuracy for multi-label classification of apparel style and sematic attributes (color, texture), thereby automating image labeling system, reducing labeling time by 3x and enhancing product recommendations by 20%.

TECHNICAL SKILLS

Languages: Python, SQL, R, HTML
Libraries: OpenCV, Keras, NLTK, scikit-learn, pandas
Tools: Git, Tableau, MS SQL Server, Azure ML Studio, Visual Studio, LaTeX, JIRA

PROJECTS

Data Warehousing for a Retail Chain *Feb 2020 – Mar 2020*
- Built BI system to analyze sales and targets of a company and provide recommendations using Tableau to increase profit margin.
- Performed ETL: created star schema for dimensional data modeling, staging tables, dimensions and fact loads using SSIS

Boston Consulting Group - Churn Prediction *Oct 2019 – Nov 2019*
- Processed raw data, cleaned negative and null values, analyzed variable distributions, handled class-imbalance and outliers.
- Used logistic regression to find significant predictors, chi-squared test to find correlation of categorical variables.
- Implemented random forest with relevant data features to classify customer churn of a power company with 94% accuracy.

Predicting Prices for Airbnb *Feb 2019 – Mar 2019*
- Carried out data preprocessing - changed numerical predictors to factors, log transformed target variable, etc; exploratory data analysis and feature engineering. Used Ridge and Lasso regression for regularization as well as feature selection.
- Applied regression models - linear regression, elastic net, and random forest models to isolate the most important predictors and predict prices of the Airbnb listings. Evaluated the performance of each model based on Mean Square Error on the test data.