

Dikshya Mohanty

Portfolio LinkedIn Github

Email: dikshyamohanty94@gmail.com Mobile: +1-704-626-9420

EDUCATION

- **University of North Carolina** Charlotte, NC
Master of Science in Data Science and Business Analytics; GPA: 3.9/4.0 Aug. 2018 – Dec. 2019
 - **Teaching Assistantship:** Applied ML (Fall 2019), Business Intelligence Analytics (Spring 2019), Big Data for Competitive Advantage (Spring 2019)
- **Indira Gandhi Institute of Technology** Talcher City, Odisha, India
Bachelor of Technology in Computer Science; Major GPA: 8.57/10.0 Aug. 2012 – July. 2016

EXPERIENCE

- **Ernst and Young LLP** Chicago, IL
Manager - Senior Data Scientist January 2020 - Continuing
 - Building an inventory management system for multilingual data (Product2Vec) for a client using deep learning.
 - Leading the CoE AI Tax team, building Machine Learning solutions for various clients in EY, both independently and in a supervisory role, leading the Staff Data Scientists on my team. Helped clients save Millions in Tax Returns and reduce the workload as well as the manpower needed by introducing and implementing novel ML solutions for Tax related problems. Actively involved in Identifying use-cases, handling communications with clients, strategising for client-specific solutions and presenting findings from experimentation using A/B testing.
 - Developing an end-to-end machine learning solution for a property tax use case focused on getting peer sets for assessing the net worth of a property better through clustering.
 - Designing an anomaly detection system to detect and extract any tax exemption qualified entities in RD credit solutions by scraping unstructured data from IRS website to categorize client's tax data into official tax categories.
 - Designing and implementing Tax chatbots using BERT.
 - Conducted research and experimentation on building a deep learning framework to extract domain-specific vectors or building a classification model, building customer NER models for extracting dates which don't have a set pattern.
 - Leading the initiative of creating the Alwin Toolkit process – panning out the overall architecture, executing, optimizing code to make our approach to the ML problems efficient and robust, in terms of resources and streamlining the entire process using Azure DevOps and Python.
- **Amazon.com** Seattle, WA
Applied Science Intern, Seller Lending Summer 2019
 - Analyzed the impact of customer purchasing trends to answer questions on seller's future success, loan payment behaviors by developing cascade models to identify 'hidden' influencers and to eventually build a seller customer base.
 - Evaluation of these newly added features was done through feature importance testing using trained XGboost, random forest models and by analyzing its p-value with regularized regression which resulted in these features being part of the top 10 percent of the existing 30k features.
 - Built an evaluation pipeline to test if addition of new features improves the current model using A/B testing. This pipeline can be used for testing any new features that the team considers useful for their model.
- **Tata Consultancy Services** Pune, MH
Data Science Developer July 2016 - July 2018

- Understanding the client's requirements, business challenges and their data, charting out the problem, defining a scope, executing the Proof-of-concept in a couple of months and delivering it to the client in a detailed report and a presentation, after which they would decide if they would want us to scale it further.
- Researched and designed text-to-vector models which helped in language understanding and generation applications such as building appropriate domain-specific vocabulary, extracting named-entities, co-reference resolution, and other language tasks.
- Designed and implemented an end-to-end ML workflow which could tackle language complexities to extract and understand patient's information to develop an online doctor which could understand patient's ailment based on varied features.
- Designed the internal workflow to process unstructured patient information extracted from PDFs using OCR to organise and segment the data into structured tables using PostgreSQL. Also stored and designed schema in MongoDB.
- Worked on designing and creating APIs to extract data from sources like PubMed and Wikipedia for domain-specific language models.

Indian Institute of Technology

Kharagpur, WB

Research Intern, Linguistics Department

Summer 2015

- **Sentiment analysis:** Designed and implemented an information retrieval and classification system for predicting user related information on Twitter. Crawled tweets from Twitter API, and extracted JSON responses using Requests module. Cleaned, parsed and segmented tweets content; determined if the user was intoxicated or not, based on the semantics of the tweet.
- **Text summarization:** Developed and implemented extractive summarization algorithms from scratch, and tested them on Wikipedia articles.

PROGRAMMING SKILLS

- **Programming Languages and Tools :** Python, SQL, R, SAS, Scala, Stata, SPSS, Excel, Tableau, PowerBI, Alteryx, Tensorflow, PySpark, Pytorch, Keras, MapReduce
- **End-to-End ML Frameworks :** Microsoft Azure Machine Learning, Amazon Sagemaker
- **Machine Learning :** Classical ML (Regression, Classification, Tree-based Models, Dimensionality Reduction Methods, Optimization techniques, Feature engineering)
- **Deep Learning :** Language Models, Deep Neural Networks (RNN, CNN, LSTM, Encoders, Transformers)
- **Statistical Methods :** Panel data Analysis, Hypothesis testing, Statistical Inferences, A/B Testing, Probability, Calculus

AWARDS

- **Ovation Award, EY FY2022:** Won an Ovation award, awarded by Americas Technology Leader. This award is for top 5% performers of all employees at EY
- **AI Challenge, EY FY2020:** Awarded first place for developing a solution to draw insights from Covid data challenge on Kaggle using FastAI
- **Bravo Awards, EY FY2020-FY2022:** Won Bravo Awards for exceptional performance on projects in every quarter of FY20, FY21 and FY22
- **Dean's List, UNCC - 2019:** Among the top 1% of my class at School of Data Science, Batch of 2019
- **Best Intern award, IIT Kgp - 2015:** Was awarded the best intern award along with a cash prize