Nikita Bhiva Raut

(224)-616-7446 | nikita.raut@u.northwestern.edu | in/nikitaraut | rautnikita77.github.io | Evanston, IL

EDUCATION

Northwestern University **Evanston, IL**

Master of Science in Computer Science | GPA: 4.0|4.0 **Expected: Mar '21**

- **Coursework**: Data Science Seminar, Advanced Deep Learning, Introduction to Artificial Intelligence, Statistical Language Modeling, Machine Learning, Natural Language Processing, Design and Analysis of Algorithms, Geospatial Vision

University of Mumbai **Mumbai, India**

Bachelor of Technology in Computer Science | GPA: 3.9|4.0 **Aug '15 - Jun '19**

- **Coursework**: Data Warehousing and Mining, Neural Networks, Advanced Databases, Data Structures, Statistics

WORK EXPERIENCE

Integrated Design Automation Laboratory, Evanston IL – Researcher **Jun '20 – Present**

- Mined product information from online shopping platforms with Selenium, BeautifulSoup and OCR to identify **200+** key product attributes driving customers' consideration-then-choice purchase decisions.
- Headed research on aspect-based sentiment mining and topic-modeling by augmenting POS tagging and Dependency Tree Parsing to detect product association relations, customer sentiments, demographics, and social influential factors from customer reviews.
- Built a full stack web survey for customer preference elicitation and worked cross-departmentally to refine the survey design.
- Designed & created a relational PostgreSQL database to support the frontend survey and preserve product information.

Northwestern University, Evanston IL | Prof. Doug Downey – Graduate Research Assistant **Apr '20 – Oct '20**

- Constructed a large corpus of real and fake biographies by scraping **1M+** existent biographies from Wikipedia and fine-tuning GPT2 transformer model to generate **50K+** fakes.
- Developed a fakes detection system to defend against fake biographies posted on social media by training LSTM and multi-channel CNN architectures and optimized embedding space initialization with text vectorization techniques like Glove and Word2Vec.

University of Mumbai, India – Deep Learning Intern **Dec '17 - May '18**

- Implemented a Real Time Face Recognition model using one-shot learning, enabling learning a class from a single labelled image.
- Extracted 68 facial landmarks with Dlib library and performed Transfer Learning with ResNet to boost accuracy by **19%.**

Computer Help, India – Machine Learning Intern **May '17 - Jul '17**

- Devised a CNN-based Damage Detection System, aimed at early detection of cracks to concrete surfaces in inaccessible places.
- Replaced slower subjective traditional human inspection procedures with fast automated processes to attain **72%** speed up.

SKILLS

Programming Languages: Python, Java, R, C/C++, MATLAB, HTML, CSS, JavaScript
Databases and Tools: PostgreSQL, MySQL, MongoDB, Tableau, Databricks, AWS, Git & Version Control
Libraries: PyTorch, TensorFlow, SciPy, NLTK, NumPy, OpenCV, Pandas, Scikit-learn, Flask

PROJECTS

Self-supervised 3D Anthropometric Measurements **Apr – Jun '20**

- Acquired clothing measurements from images of client clicked at four angles and SMPL point cloud with errors of up to **40 - 50 mm**.
- Deformed a pre-defined body mesh by comparing its projection to input views by training Siamese network on adversarial loss.
- Measured circumference paths for 18 body parts on fitted 3D model and provided final estimates with 1D embeddings.

COVID-19 spread prediction using Graph Neural Networks **Feb - Mar '20**

- Formulated a highly scalable pipeline to analyze spread of COVID 19 in United States by exploiting SageConv and Message Passing.
- Modeled a network having census and travel data as links, provinces as nodes and JHU data as features to predict cases.

Image Inpainting for restoring corrupted regions of Image **Jan - Mar '20**

- Proposed a novel method for restoring corrupted regions of an image by passing its k-nearest pixels through a BiLSTM network.
- Revamped the existing model by overcoming some of its shortcomings like shape dependency, computational expense, unused pixels, discontinuity in processing order and perception quality that helped increase the SSIM score by **13%.**
- Computed the pixels in final image by solving Poisson equations to preserve intensities between restored and original regions.

Text-2-SQL sequence-to-sequence translation **Jan - Mar '20**

- Addressed the cross-domain Text-2-SQL generation task by implementing a syntax tree-based recursive encoder-decoder model.
- Exploited structural format of SQL language by employing two encoders for text and schema, and eight decoders for prediction of various SQL query tokens, thus enhancing generalizability on unseen schema with an exact matching score of **88%**.
- Refactored transformer model to make them partially autoregressive by allowing for flexible sequence generation.

Data Analytics on Chicago Police Dataset **Sep – Dec '19**

- Gained insights on behavior of officers charged with misconduct by analyzing Chicago Police Dataset and **1000+** filed complaint reports by leveraging relational analytics, graph analytics, machine learning, time series analysis and NLP models.
- Presented findings on officer misconduct and misuse of power based on compelling visualizations in Tableau and D3.js.