KEDAR DABHADKAR

dkedar@cmu.edu | (734) 819-0242 | linked.com/in/dkedar7 | dkedar7.github.io

EDUCATION

Carnegie Mellon University | GPA: 3.65/4.0 — **Pittsburgh, PA**
Master of Science in Chemical Engineering — **Dec. 2018 (expected)**
(Specialization: **Data Analytics, Machine Learning**)

Relevant Coursework:
Process Systems Modeling (06-665), Introduction to Machine Learning (10-601), Computer Science in Chemical Engineering (06-611), Applied Data Science (16-791)*, Introduction to Deep Learning (11-785)*. **Currently enrolled*

Institute of Chemical Technology — **Mumbai, India**
Bachelor of Chemical Engineering — **May 2017**

SKILLS

Programming Languages: *Proficient*: Python, R, SQL; *Intermediate*: JAVA; *Basic*: Bash, FORTRAN, C++, HTML.
Software: Apache Hive, Apache Spark, Tableau, MATLAB, GAMS, ALAMO. **Databases:** PostgreSQL, MySQL, MSSQL.
Packages: Pandas, TensorFlow, PyTorch, Keras, scikit-learn. **Cloud Platform**: AWS EC2.

RESEARCH EXPERIENCE

Graduate Researcher — **Pittsburgh, PA**
Data-driven Modeling of Process Performance |Python|R|MATLAB|ALAMO — **Jan 2018-present**
*Was presented at **'Big Data and Process Engineering: Opportunities and Limits', Paris**, France on September 5, 2018.*

- Compared applicability of various time-series smoothing methods using the Dickey-Fuller unit root test.
- Modeled multivariate process data using ARIMA with exogenous variables (ARIMAX), ALAMO and NARX time-delayed neural networks and by employing custom-designed time-averaged RMSE and R^2 as the performance metrics.
- Laid the framework and data pipeline to replace first-principle based approach by statistical modeling.

PROJECTS

Analysis of Medical Records of Cancer Patients Using Natural Language Processing | Python — **Pittsburgh, PA**
Third Prize, Hackathon, North American Association of Central Cancer Registries (NAACCR) — **June 2018**

- Analyzed Electronic Medical Records (EMRs) of 10,000 cancer patients to classify according to cancer site.
- Scored an average F1 of 0.91 on held-out data with an ensemble of Naïve Bayes, Random Forests and SVM.

Pattern Recognition in Electroencephalogram (EEG) of the Brain | Python | MATLAB — **Pittsburgh, PA**
First Prize, Hackathon, Auton Lab, Carnegie Mellon University and Phillips — **March 2018**

- Cleaned, pre-processed noisy EEG data to induce stationarity and transformed into a sequential window matrix.
- Predicted the occurrence of Cyclic Alternating Pattern (CAP) with an accuracy of 58% using logistic regression.

Named Entity Recognition | Python | AWS — **Spring 2018**

- Built feature engineered logistic regression models to extract information from about 50,000 sentences.
- Deployed an AWS EC2 p2.xlarge instance to handle heavy computations and got an average F1 score of 0.94.

Speaker Verification | Python | PyTorch | AWS — **Fall 2018**

- Trained a Convolutional Neural Network (CNN) for speaker identification on 125 GB data, extracted embeddings, compared speakers based on cosine similarity and obtained EER (Equal Error Rate) of 13.7 %.

Part of Speech Tagging | Python — **Spring 2018**

- Trained a Hidden Markov Model (HMM) using forward-backward algorithm to tag all words from 3500 sentences with their respective parts of speech.
- Achieved a negative log likelihood of 97 on the held-out data.

Bilingual Speech Synthesis System | Python | TensorFlow | AWS — **Spring 2018**

- Developed a deep network to detect code-switching in English-Spanish corpus using data from 84 speakers.
- Classified code-switching into 5 types and hypothesized that doing this could improve speech recognition.

Time Series Analysis of Currency Valuation | Python — **Fall 2017**

- Implemented descriptive statistics, various smoothing and stationarity induction methods, and auto-correlations to analyze valuation of the Indian National Rupee against the US Dollar.
- Employed web-scraping to perform live one-day-ahead predictions with ARIMA (MSE=0.05) and LSTM (MSE=0.03).

AWARDS AND LEADERSHIP

Narotam Sekhsaria Foundation Postgraduate Scholarship — **June 2017**
Treasurer, General Student Body, Institute of Chemical Technology — **July 2015-July 2016**
Event coordinator, ICT Marathon — **Dec 2015**