

KEDAR DABHADKAR

dkedar@cmu.edu

(734) 819-0242

[linked.com/in/dkedar7](https://www.linkedin.com/in/dkedar7)

Portfolio: [dkedar7.github.io](https://github.com/dkedar7)



EDUCATION

Carnegie Mellon University | GPA: 3.65/4.0

Master of Science in Chemical Engineering

(Specialization: **Data-Driven Decision-Making**)

Pittsburgh, PA

Dec 2018 (expected)

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).

Institute of Chemical Technology

Bachelor of Chemical Engineering

Mumbai, India

May 2017

SKILLS

Programming Languages: *Proficient:* Python, R, SQL.

Intermediate: JAVA.

Basic: Bash, FORTRAN, C++, HTML.

Software: MATLAB, GAMS, ALAMO, Tableau.

Databases: PostgreSQL, MySQL, MSSQL, Oracle, MongoDB.

Packages: Pandas, TensorFlow, PyTorch, Keras, scikit-learn.

EXPERIENCE

Graduate Researcher

Data-driven Modeling of Process Performance |Python|R|MATLAB|ALAMO

Master's Research, Carnegie Mellon University in collaboration with Air Liquide, Germany

Presented at 'Big Data and Process Engineering: Opportunities and Limits', Paris, France on September 5, 2018.

- Documented literature to model industrial reactors as hybrids of data-driven methods and first-principles.
- Modeled multivariate process data using ARIMA with exogenous variables and ALAMO.
- Currently implementing response variable-constrained autoregressive neural networks with exogenous inputs.

Pittsburgh, PA

Jan 2018-present

PROJECTS

Analysis of Medical Records of Cancer Patients Using Natural Language Processing | Python

Third Prize, Hackathon, North American Association of Central Cancer Registries (NAACCR)

Pittsburgh, PA

June 2018

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

Detection of Patterns in Electroencephalogram (EEG) of the Sleeping Brain |Python|MATLAB

First Prize, Hackathon, Auton Lab, Carnegie Mellon University and Phillips

Pittsburgh, PA

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%.

Named Entity Recognition |Python|AWS

Spring 2018

- Built logistic regression models to extract information from about 50,000 sentences with subsequent feature modifications (average F1 score of 0.94).
- Deployed an AWS EC2 p2.xlarge instance to handle heavy computations.

Data-driven Process Scheduling |GAMS

Spring 2018

- Improved operation of a manufacturing facility using process data to find optimal schedule.
- Decreased cost of operation by 28% using CPLEX (MILP) solver on GAMS.

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.
- Got a negative log likelihood of 97 on the held-out data.

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

NS Foundation Postgraduate Scholarship

June 2017

Treasurer, General Student Body, Institute of Chemical Technology

July '15- July '16

Best Technical Presentation, Indian Institute of Chemical Engineers

Feb 2016

Event coordinator, ICT Marathon

Dec 2015