# Nikhil Tilak

**Website**: https://nikhiltilak.github.io/

**Email** nikhiltilak1991@gmail.com
**Phone** +1 (732)-500-6319
**GitHub** github.com/NikhilTilak
**LinkedIn** linkedin.com/in/nikhil-tilak

## Selected Data Projects

***Bookend*** GitHub, video

- Trained an ensemble classifier model on books scraped from project Gutenberg to predict the authorship of a snippet of text with 93% accuracy.
- Led a team of four and was responsible for dividing tasks and establishing a GitHub-based workflow to maximize productivity.
- Implemented a bag-of-words model which gave the highest prediction accuracy score (85%) among the models in the ensemble.
- Project placed 1st in the final project competition of Erdos data bootcamp.

***Sudoku-Solver*** Github, App

- Created an application to solve a Sudoku puzzle correctly given its image as input.
- Wrote a custom pipeline which processes the image, identifies the filled digits using OCR and a neural network and produces a solution.
- Deployed a Docker containerized Dash/Plotly app to Google Cloud (GCP).

***IMDB Movie review sentiment analysis*** Kaggle kernel

- Trained a neural network to predict if a given movie review is “positive” or “negative”.
- Used a word2vec model trained on the reviews to generate semantic word embeddings.

***BreweryXplorer*** GitHub, Dashboard

- Designed and deployed an interactive web app that lets users browse and search 3000+ breweries and pubs in the United States. Scraped and cleaned unstructured brewery data gathered from Wikipedia and other web sources.
- Designed an interactive Dashboard using Dash/Plotly which was deployed to Heroku.

## Skills

**Programming/Scripting:** Python, C/C++, SQL, Excel, MATLAB

**Packages:** NumPy, SciPy, Pandas, Scikit-Learn, OpenCV, Natural Language Toolkit, TensorFlow

**Cloud and related tools:** Google Cloud Platform (GCP), Heroku, Docker.

**Machine learning and Statistics:** Data scraping and wrangling, dashboards, machine learning models for classification, regression, clustering, outlier detection and forecasting. Natural language processing (NLP), Convolutional Neural Networks (CNNs), embeddings, sentiment analysis, hypothesis testing, A/B testing.

## Professional Experience

**PhD Candidate, Physics** | Rutgers University, New Jersey, USA | (2015-present)

- Designed and performed state-of-the-art experiments to explore electronic properties of twisted two-dimensional materials using Scanning Tunneling Microscopy (STM).
- Analyzed multidimensional experimental data using Python to extract weak signals from noisy data.
- Used clustering algorithms for feature identification and background subtraction on STM images.
- Expert in breaking down complex ideas and presenting them to technical & non-technical stakeholders.
- Extensive writing experience which led to four high-impact publications in peer-reviewed journals.

## Data Science Certificates

**Erdos Institute Data Science Bootcamp** (May-July 2020, 1st place winning final project.)

**Coursera:** Applied Data Science with Python Specialization (four separate courses) (Jan-Jun 2020).

## Education

PhD in Physics & Astronomy, Rutgers University, NJ, USA. | 2015-Feb 2023 (expected)
B. Tech. (Engineering Physics), Indian Institute of Technology, Guwahati, India. | 2010-2014