# PHILLIP CHAU

(408)-706-8868 | pchau@seas.upenn.edu| https://phillipchau.github.io/ | github.com/phillipchau | US Citizen

## Education

**University of Pennsylvania | School of Engineering and Applied Science |** Philadelphia, PA *May 2023*

- *Candidate for **Bachelor of Science** in Computer Science*
  - *Cumulative GPA:* 3.92/4.00 | *Minor:* Statistics, Mathematics, Data Science
  - *Relevant Coursework:* Scalable and Cloud Computing, Data Structures and Algorithms, Probability, Python for Data Science, Statistical Inference
- *Candidate for **Master of Science Engineering** in Data Science*
  - *Cumulative GPA:* 4.00/4.00
  - *Relevant Coursework:* Big Data Analytics, Machine Learning, Database and Information Systems

## Experience

**NASA | Software Engineering Intern |** Mountain View, CA *June 2020 – December 2020*

- Developed full-stack platform with React, Express.js, Node.js, MySQL and DynamoDB to manage real time aerial vehicles
- Developed frontend framework to simulate and track flight plan mis-compliance, visualizing results with Google Maps API

**STEM-Away | Machine Learning Intern |** Santa Clara, CA *June 2020 – July 2020*

- Created web crawlers through BeautifulSoup to extract data (36k+ entries) from Discourse Forums
- Performed data cleaning, analysis, pre-processing and feature engineering through Python, Pandas, Numpy
- Developed forum classifier using BERT, TF-IDF, Word2Vec NLP models to predict associated forum for posts

**Wharton Customer Analytics | Senior Analyst |** Philadelphia, PA *February 2021 – Present*

- Created Time Series Forecasting and Econometric Models using 100,000+ consumer data entries to predict impact of social marketing strategies on client's revenue streams and brand affinity and presented recommendations to client

**University of Pennsylvania | Teaching Assistant |** Philadelphia, PA August *2021 – Present*

- TA for CIS545: Big Data Analytics

## Projects

**PennBook | Personal Project |** Philadelphia, PA *October 2020 – December 2020*

- Created secure and scalable Facebook clone with features including user profiles, chats, content feeds
- Built with Node.js, JQuery, Express.js, AWS DynamoDB and hosted on AWS EC2
- Designed scalable news recommendation algorithm and news feed page using Apache Spark

**Crash Prediction Model | Personal Project |** Philadelphia, PA *November 2020 – December 2020*

- Analyzed Pennsylvania Vehicle Crash Data over a decade through EDA with Python, Pandas, Seaborn
- Created classification models through Random Forest, Logistic Regression to predict fatal crashes based on driving-related factors

**Fake News Analysis | Personal Project |** Philadelphia, PA *April 2021 – May 2021*

- Analyzed 10,000+ tweets about 2016 US Election through EDA with Python, Pandas, Plotly
- Cleaned and tokenized text and conducted sentiment analysis to understand factors that contribute to spread of fake news
- Created topic model of fake news through LDA and trained Naïve Bayes and Random Forest models to classify fake news

**COVID Dashboard | Personal Project |** Philadelphia, PA *March 2021 – May 2021*

- Created visualization tool to analyze how socioeconomic factors impacted Covid cases and deaths across America over time
- Integrated NYT and Twitter API for real time news and enabled user authentication via Facebook and Google
- Built with React, Express.js, SQL, Firebase, AWS DynamoDB

## Skills

**General**: Java, Python, JavaScript, Bash, Git
**Data Science**: Pandas, Numpy, Plotly, Scikit-learn, NLTK, Apache Spark, SQL, AWS DynamoDB, AWS S3, MongoDB
**Web:** React, HTML, CSS, EJS, Node.js, Express.js
**Additional**: Problem Solving, Innovative, Committed, Patient, Articulate, Collaborate, Agile