

✉ mindyng88552@gmail.com
🌐 mindyng.github.io
☎ 5105082455
in
www.linkedin.com/in/mindyng85
🔗 mindyng
I empathize with customers to relentlessly meet their needs by being a collaborative thought-partner to stakeholders.

I hunger for business' growth through proactive exploration of ambiguous, complex problems and providing creative data recommendations to inform strategy.

I am passionate about user journeys, product innovation and defining success metrics. I am a data journalist to diverse disciplines and levels telling the business story of how a product is solving real issues.

I love data modeling, visualization, experimentation, statistical methods, forecasting, optimizing processes, teaching and managing up.

Skills

PROJECT MANAGEMENT

Defining KPI's
Scoping Out Business Problem
Defining Metrics of Success
Cross-Functional Collaboration
Insights to Stakeholders
Results/Output-Driven

LANGUAGES

SQL
Python

DATA ENGINEERING

Google Cloud Platform
PostgreSQL
Snowflake
dbt
Meltano
Dataform
Prompt Engineering

DATA WRANGLING

Data Cleaning
Data Integrity Checks / Assertions

STATISTICS

Descriptive Statistics
Inferential Statistics
Experimentation
Hypothesis Testing
A/B Testing

MODELS / MACHINE LEARNING

Linear Regression
Logistic Regression
Natural Language Processing (NLP)
Large Language Models (LLMs)

BUSINESS ANALYTICS

User Journeys
LTV Analysis
Churn Prediction
Cohort Retention Analysis
Time Series Analysis

VISUALIZATION / BUSINESS INTELLIGENCE

Looker
Tableau
Superset

MINDY NG

DATA ANALYST

Employment

Upwork/Fiverr
Freelancer
Data Analysis/Science and AI Applications

Remote
Oct. 2022 to Current

Mercari US
Data Scientist

Palo Alto, CA
Sept. 2022 to Oct. 2022

- Helped build out search team's new metrics table in Dataform to power dashboards and assess experiment results.
- Analyzed price elasticity between highest GMV/most searched for category groups to understand buyers' demand change in relation to item price changes.

Mercari US
Business Intelligence Analyst

Palo Alto, CA
Aug. 2021 to Aug. 2022

- Created Search RFM segmentation to personalize search and increase GMV.
- Investigated query chaining to understand search engine performance as well as searchers' persistence for casual and power users' high value purchases.
- Performed user journey analysis across web to provide insights on platform's highest touch points for ML efforts to converge on.
- Analyzed navigation vs search activity to improve our UI to cover all our users' intent to drive north star metrics.
- Built statistical significance framework for experimentation at scale to add criteria for product roll out.

Forethought
Implementation Engineer

San Francisco, CA
July 2020 to Sept. 2020

On the Customer Experience team, leading all technical requirements and touching all aspects of the business: Engineering, Product, Sales and Customer Success

Implemented: State-of-the-art NLP models to help clients be geniuses at their job
Involved: Data Engineering, Data Science, Machine Learning/Artificial Intelligence, Business Intelligence -- owning whole data pipeline Post-Sale

- Queried MongoDB to create customer business rules.
- Designed AI Training datasets to feed into XLNet and BERT models using Jupyter Python notebooks.
- Analyzed trained models' performance to deploy best automated NLU models for clients.
- Verified live models' predictions were successful via API calls to clients' Salesforce Help Desks.
- Reduced client's SPAM from 64% to less than 1%.
- Helped save client >\$20,000 in human labor cost from Customer Support Agents manually labeling tickets.
- Completed data analysis that contributed to signing of >\$400,000 deal with Instacart.

Immuno Concepts
Quality Control Analyst

Sacramento, CA
July 2010 to Apr. 2019

- Built linear regression models to determine whether or not products were drifting from quality.
- Tracked trends and outliers to make manufacturing recommendations to management to create efficiencies and increase profit margins.
- Created product performance reports to drive key business investments for following quarter.

University of California, Davis
Research Associate

Davis, CA
Jan. 2005 to Dec. 2008

- Through repeated experimentation explored sigma70 subunit architecture to characterize macromolecular complexes involved in transcription of growth-related genes.
- Narrowed down which protein chain substitution in antibody-derived proteins fit best with research aims in pre-targeting radioimmunotherapy for Non-Hodgkin's Lymphoma.

Projects

Sentiment Analysis on LLM Chatbots
Examined popular chatbots' user sentiment to help inform product roadmap and brand strategy.
Nov. 2023 to Dec. 2023

LTV and Churn Analysis on Lyft Drivers
Examined supply and demand of ride-sharing business by looking at trends and contributions to LTV and churn. Also performed retention cohort analysis to inform business strategy for two-sided marketplace.
Sept. 2023 to Sept. 2023

Recommendations to Improve Skip Rates for Meal Subscription Business
Used customer facts and order history to provide data-backed advice to remedy increased YoY skip rates for Tovola.
July 2023 to July 2023

Time Series Forecasting on Uber Eats' Vendors
Utilized 7,911 samples of date-stamped data and predicted which vendors were worth continuing business with based on ROI.
Dec. 2018 to Dec. 2018

Trended each vendors' data with Facebook's Prophet. Trends performed over a span of 15 months. Data further broken down into weekly and daily trends. Resulting model performance based on 30-day horizon producing 0.01 - 0.03 RMSE.

Publication

American Chemical Society Publications -
Mapping Protein-Protein Interactions by Localized Oxidation: Consequences of the Reach of Hydroxyl Radical
Results provided fundamental information for interpreting protein foot-printing experiments in other systems.
Feb. 2009 to Apr. 2009

Volunteering

CoronaWhy
Machine Learning Engineer
CoronaWhy is a globally distributed, volunteer-powered research organization of 1000+ members. They are using DS and AI to assist the medical community and policy makers answer key questions related to COVID-19. It's supported by Google, Amazon, NASA and other companies.
Apr. 2020 to June 2020

I was embedded within the Vaccine/Therapeutics Task team, helping the Paper Study Classification group build baseline models to filter papers based on study design.

Education

Springboard, Data Science Career Track
Jan. 2017 to Dec. 2017

University of California, Davis
Genetics Bachelor's of Science
Sept. 2003 to Dec. 2007