

Justin Chen

chen.justi@northeastern.edu • 650-388-0985 • <https://jyhchen.github.io/>

Availability: July/August 2023

EDUCATION

Northeastern University, Khourey College of Computer Sciences

Candidate for Bachelor of Science in Data Science and Business Administration

Awards and Honors: Dean's Scholarship, Dean's List, University Honors Program

September 2019 - Present

Expected: August 2023

GPA: 3.51/4.0

Coursework In: Object Oriented Design; Database Design; Machine Learning and Data Mining; Data Visualization; Sports Media, Marketing, and Communications; Financial Data Analytics

Activities: Baseball Student Manager, WRBB Sports, Taiwanese American Student Association, Intramural Sports

SKILLS AND INTERESTS

Computer Languages: Java, Python, Racket, SQL, HTML, CSS, Javascript, R

Frameworks and Packages: Django, React.js, D3.js, SQL Server, AWS

Software: Git, IntelliJ, Jupyter, Visual Studio Code, Tableau

Interests: Baseball, Basketball, Sports Collectables, Movie Soundtracks, Public Transportation, Biking

Other: Social Media, Sports Writing and Broadcasting, Chinese (native speaker)

WORK EXPERIENCE

Web Analytics and A/B Testing Co-op

Harvard Business Review

Boston, MA

July 2022- Present

- Designed and ran A/B tests to analyze consumers responding to changes in design on the HBR site
- Created interactive visualizations and dashboards with R and Adobe Analytics to summarize experiment findings
- Automate workspace backup and creation with R and Python, utilizing Adobe APIs
- Explored consumer behavior in article and podcast recommendations, subscriber flow, and newsletter engagement

Student Manager and Analyst

Northeastern University Baseball

Boston, MA

January 2020- Present

- Collected batted ball and pitching metrics with Rapsodo and HitTrax
- Charted pitch-by-pitch game data with Excel and visualized trends and statistics with Python and Powerpoint
- Created scouting reports and pitch design using Synergy
- Analyzed player video for pitch tipping and mechanical inconsistencies

Baseball Systems Development Co-op

Baltimore Orioles

Baltimore, MD

July 2021- January 2022

- Rebuilt data pipelines for player stat views on the team's internal site with Django
- Developed amateur player evaluation forms and tables for scouts with React
- Made quality checks and internal ID mapping for imported third-party data and created Slack notification bots
- Automated web scrapers with AWS lambda functions

PROJECTS

NCAA Basketball Prospect Grader (Python/Jupyter Notebook)

- Built models to rank NBA Draft prospects and project their NBA win shares based on their collegiate careers
- Collected dataset of 15,000 NCAA players with over 45 features and their NBA contributions with data scraping
- Used one-hot encoding and recursive feature elimination find the optimal feature subset
- Trained regression models with cross-validation, including Random Forest, AdaBoost, and Multilayer Perceptron
- Ranked players for the 2022 Draft, with results similar to ESPN's projections

MLB Repository (MySQL/Python)

- Created a relational database to store biographies and statistics for players and teams
- Implemented functions and triggers to update tables with additions or deletions of records
- Used Python to create a command-line application to access player and team information and statistics

Shot Grader: an Alternative to xG (WHKYHAC 2022) (Python/Jupyter Notebook)

- Graded PHL and Olympic shots based on shot destination results, different than the two-true outcomes of expected goals
- Explored importance of skater traffic and shot type when analyzing the quality of a shot
- Discussed data constraints in the realm of women's hockey