

Contact

www.linkedin.com/in/yaoya0
(LinkedIn)
ctools.umich.edu/osp-presentation-tool/viewPresentation.osp
(Personal)

Top Skills

Machine Learning
Python
Statistical Data Analysis

Languages

German (Professional Working)
English (Native or Bilingual)
Chinese (Limited Working)

Certifications

Certificate of Completion: Python Course
edX Verified Certificate for Programming with Python for Data Science
edX Verified Certificate for Implementing Predictive Analytics with Spark in Azure HDInsight
edX Verified Certificate for Programming in R for Data Science
edX Verified Certificate for Applied Machine Learning

Honors-Awards

Second Place in Best Capstone Poster at SMU
Tech-Arb Challenge

Yao Yao

Data Scientist at Kanopy
Orange County, California Area

Summary

American Citizen

Interests and skills: Python, R, Databricks/Spark, Multiprocessing, Azure, Computer Vision, C++, Keras, Tensorflow, Processing 3, Machine Learning, Statistical Modeling, Time Series, Predictive Analytics, Visualizations

Team leader, organizer, writer, and presenter for narrating dataset, problem, workflow methods, interpretations, and applications

Mastered: data scraping, proportional sampling, imputation, ARIMA, clustering, Naive Bayes, Stanford NLP, decision trees, random forest, XGBoost, SVM, neural networks, Hidden Markov, OOP, polymorphism, feature importance, reduction, LDA, PCA, visualization, optimization, classification, validation, regression, Gaussian mixtures, audio separation

If you are new to and/or would like to get into the field of data science, I did some presentations about it:

<https://www.youtube.com/watch?v=Fiz1Tn7ogP4>

https://www.youtube.com/watch?v=_ZeTP4wpDSU

Experience

Kanopy
Data Scientist
October 2019 - Present (6 months)
Orange County, California Area

Recommendation Engine

Price Optimization

Lifetime Value

Download the Kanopy app and link it with your library card to stream 10 films per month for free

Viral Launch

Data Scientist

November 2018 - October 2019 (1 year)

Indianapolis, Indiana Area

- Soloed 6+ projects from data orchestration, workflow design, to production code for online software release
- Used local and Azure cloud multiprocessing to forecast time series predictions for 50+ million search terms
- Optimized key features for ad campaigns to generate best ROI for ad bid, ad budget, and sales margins
- Used feature importance to find top search terms that generated most revenue for top 20+ million products
- Applied computer vision and split testing to optimize product pictures to generate best sales conversion

Southern Methodist University

Graduate Student Researcher

January 2017 - August 2018 (1 year 8 months)

Dallas/Fort Worth Area

- 11 Major Projects + Capstone Graduation Project
- Team leader, writer, and presenter for narrating dataset, problem, workflow methods, interpretations, and applications
- Yelp's Review Filtering Algorithm (Thesis, Published), ML Audio Source Separation, Taught Databricks & Blockchain
- Mastered: data scraping, proportional sampling, imputation, ARIMA, clustering, Naive Bayes, Stanford NLP, decision trees, random forest, XGBoost, SVM, neural networks, Hidden Markov, OOP, polymorphism, feature importance, reduction, LDA, PCA, visualization, optimization, classification, validation, regression, Gaussian mixtures, and audio separation
- Check Project drop down for more info
- Second Place in Best Capstone Poster at SMU

*Research

Graduate NLP and Semantics Researcher

March 2016 - December 2017 (1 year 10 months)

Greater Chicago Area

- Researched new product and market creation by analyzing large scale, online interactions on Reddit
- Parsed unstructured data into relational databases by applying NLP and semantic tagging via Python
- Analyzed and tabulated data accumulated from various APIs via SQL databases

U. S. Steel

Product Technology Management Associate

September 2012 - December 2014 (2 years 4 months)

Greater Detroit Area

- Automated multiple processes that took many hours on the computer and condensed them into a few seconds
- Wrote macro programs in VBA and designed Excel templates for steel qualifications to be automated and formatted consistently, saving time and increasing the productivity of customer technical service.
- I have written numerous visual basic codes that dealt with generating graphs, control charts, and summary reports for lab data and manufacturing and shipping logistics
- Created a form in Java for data entry at a manufacturing plant to keep track of scraps, rejects, and the costs involved. The form automatically generated a monthly report in Excel. Taught people how to use my system
- Developed new geometries for existing parts using CAE
- Used Hypermesh software to compare and change the geometries of OEM parts by comparing steel properties to that of aluminum. Used Hyperform to create a binder and addendum to investigate the formability of the part by using the forming limit diagram, the effects of gravity, and residual strain.
- Familiar with LS-PrePost and LS-Dyna (Linux) software for more FEA rendering.
- Took advanced metal forming, die troubleshooting, and supplemental FEA classes at Oakland University in collaboration with Chrysler LLC with instructor Stu Keeler among others in the auto industry

Organic Photovoltaic Cells: Polymer Solar Cell Synthesis and Application

Undergraduate Research with Professor Kim, MSE Department
September 2010 - May 2012 (1 year 9 months)

Ann Arbor, MI

- Meticulous with stoichiometry measurements, reaction process, and extraction procedures for maximum product yield.
- Characterized the polythiophene with spectroscopy methods such as NMR, IR, and C-NMR
- Created small scale polymer solar cells and applied protective arachidic acid using spin coating methods
- Responsible for photoluminescence, UV visibility, chromatography, and voltage testing
- Took Quantum Mechanics and Polymers to be part of project. Focused on p-n junctions, band structure, molecular orbital levels, and exciton triplet states pertaining to the application of solar cells.

Results withheld; graduate student I collaborated with has yet to do his PhD thesis/dissertation defense

Residence Hall Information Technology Central Services (ITCS) IT Staff

June 2008 - May 2012 (4 years)

Ann Arbor, MI

- Fixed 100+ infected computers. Installed, updated, and maintained software on 400+ computers.
- Advocate of internet safety and virus prevention, which made computers run efficiently

M-Wrap (Warmilu), Warming Blanket for Premature Infants

Undergraduate Senior Design with Professor Shtein and Professor Thornton, MSE Department

September 2011 - February 2012 (6 months)

Ann Arbor, MI

- Designed 3D blanket model using finite element analysis with Comsol to adequately warm a premature infant and created dimensions of insulating material to keep the infant warm for 4 hours.
- Executed thermodynamic and kinetics simulations of exothermic supersaturated sodium acetate.

- Fabricated a working prototype based on computer model dimensions and ran tests with LabView with thermocouples to see heat transfer and retention rates.
- Developed first generation Warmilu blanket that complies with medical and safety regulations in accord with Mott's children's hospital.

Behavioral Game Theory on eBay: Support Vector Application (Published)

Undergraduate Research with Professor Saigal, IOE department
September 2009 - August 2010 (1 year)

Ann Arbor, MI

- Derived Nash Equilibrium equations to model how much buyers should bid based on the perceived value of item.
- Wrote a program that collected user data from eBay's API system.
- Imported data via Matlab into a multidimensional analysis tool (Support Vector Machine) that created hyperplanes to formulate seller reliability index based on multiple seller attributes.
- My results: 22% of the transactions on eBay are fraudulent. 97% seller rating equals 70% seller reliability.
- To prevent fraudulent transactions from occurring, I outlined a form of 3rd party insurance which held the seller as well as the buyer responsible for the transaction and making sure that the product is genuine, which omits the refund process, thus saving the shipping fee and delivery time.
- Took Microeconomics, Programming and Data Structures to be a part of project. Focused on algorithm development for structured data types, pointers, linked data structures, abstract data types, recursion, arrays, and parameter passing methods.

Education

Southern Methodist University

Master of Science - MS, Data Science · (2017 - 2018)

University of Michigan

B.S.E, Material Science Engineering · (2007 - June 2012)

Oakland University

Master of Business Administration - MBA, non degree courses · (2015 - 2016)

Upper Arlington High School

