

Yang Gao

Curriculum Vitae

420 W. Dayton St., Apt 649
Madison, WI 53703
☎ (267) 403 6781
✉ yang.gao@wisc.edu
🌐 www.yanggao.ca
🌐 [linkedin.com/in/yang-g](https://www.linkedin.com/in/yang-g)

Education

- 2021-2025 **Ph.D. student in Operations and Information Systems, UW-Madison**
(expected)
 - Specializing in solving business problems using interpretable Natural Language Processing techniques
- 2011-2013 **M.Sc. in Chemical Engineering, University of Alberta, Boluk Lab**
 - Specializing in sustainable nanomaterials
 - Thesis: The Production and Characterization of Cellulose Nanofibrils
- 2006-2011 **B.A.Sc. in Nanotechnology Engineering, University of Waterloo**

Work Experience

- 2019-2021 **Data Science Manager, Platt Labs, Wharton Neuroscience Initiative, University of Pennsylvania, Philadelphia**
Applied machine learning and statistical analysis to advance scientific research
 - Built machine learning models and statistical analyses to detect human stress state from biosignals
 - Developed signal processing pipeline for wireless electroencephalography devices in Python & Matlab
 - Developed interpretable consumer segmentation based on natural language processing of social media data
 - Oversees computer vision project on white blood cell classification from blood smear images
 - Built central data management systems, including a PostgreSQL database to improve data integrity, and migrated to Amazon Web Services to facilitate access for international collaborators
 - Supervised over 10 people, and held workshops on machine learning and data science
- 2019-2020 **Lead Machine Learning Engineer, Omdena, Remote**
Worked in collaboration with leading NGOs on machine learning solutions for social good
 - Led a team of 10 people on a sentiment classification model to detect gang violence in social media
 - Directed ensemble model building, designed algorithms, and coaching team members on machine learning methods
 - Led a team of 5 on a project to detect Amazon forest fires using Convolutional Neural Nets
- 2016-2019 **Project Lead, Operations Transformation, Bank of Montreal, Toronto**
Led data-driven operational process improvements
 - Performed big data analytics to identify saving opportunities, increasing benefits by \$20MM
 - Led optimization projects on electronic conversion, production optimization, demand management
 - Forecasted \$100MM operational budget, and provided variance analysis to VP
- 2014-2016 **Senior Officer, CRM Database and IT, St. Michael's Hospital Foundation, Toronto**
Implemented data process automation and improved data health
 - Led a team of cross-functional resources to identify data and process gaps, improving data health and donor experience through automation and data integrity checks
 - Managed CRM database and performed segmentation based on financial and donor data
 - Managed database upgrade and online-offline database integration projects
 - Automated data entry and management processes resulting in 80% effort reduction

Academic Papers

Associating Concepts to Consumer Segments Using Social Media Data to Improve Targeting and Positioning, *Work in Progress*,

Yang Gao, Emaad Manzoor, Jonah Berger, Michael Platt

A Machine Learning Approach to Identifying Objective Biomarkers of Anxiety and Stress, *Nature Communications (Revise & Resubmit)*,

Arjun Ramakrishnan, Adam Paredes, William Lynch, Christopher Molar, Yang Gao, Dhaval Bhatt, Michael Platt

Other Projects

Computer Vision, NLP **Image Captioning using CNN and RNN**, code: <http://bit.ly/yangcnrnn>
This model combines computer vision and natural language processing to recognize objects in photos and generate captions that closely match the photos.

NLP **Robot Name Generation**, code: <http://bit.ly/yangrobot>

Character-level language model with RNN for robot name generation. Conditioned RNN on popular robot and AI names from the 19th century to latest trends, then sampled characters according to probability distribution for each character to produce new robot names.

Sequence Generation **Jazz Music Generation with LSTM**, code: <http://bit.ly/yangjazz>
Trained LSTM neural net on jazz music, and generated jazz music by sampling the most likely note at each step based on previous notes.

Computer Vision **University of Alberta Nanorobotics Team (President)**

Managed team of 9 to build a robot of ferrofluidic nanoparticles controlled by electromagnets. Autonomous mobility challenge executed with MATLAB edge detection that provided direction to electromagnets. Placed 3rd in the IEEE RAS Microrobot Microassembly challenge.

Academic Service

Ad Hoc Reviewer: *Frontiers in Neuroscience*

Skills

Computer Python, MATLAB, R, SQL, VBA, LaTeX, Google Colab, Amazon Web Services

Office Excel, Access, Word, PowerPoint, Google Suite

Languages English, Mandarin, Finnish, French