Justin A. Gould

Senior Data Scientist
The Data Mine
Purdue University
West Lafayette, IN

gould29@purdue.edu
https://gouldju1.github.io/gouldju1/
https://github.com/gouldju1
Phone: +1 (248) 877-0751

Work Experience

12/2020-Present
Senior Data Scientist
The Data Mine
Purdue University, West Lafayette

04/2020-12/2020
Product Owner, NLP Services and Connected Vehicles
Ford Motor Company
AI Advancement Center - Natural Language Processing Research
Smart Mobility Analytics - Connected Vehicles Products and Platforms

11/2019-09/2020
Machine Learning & Artificial Intelligence Scientist
Ford Motor Company
AI Advancement Center - Natural Language Processing Research

01/2019-11/2019
Data Scientist
Ford Motor Company
Smart Mobility Analytics - Driver Assistance Technology

08/2018-12/2018
Data Science Consultant (M.S. Experiential Learning)
Stryker Medical Devices
Orthopaedic Instruments

06/2018-08/2018
Data Science Intern
Ford Motor Company
Product and Sustainability Analytics

01/2018-05/2018
Data Science Consultant (M.S. Experiential Learning)
Meijer, Inc.
Marketing - Advanced Analytics

06/2016-12/2018
Data Science Intern
DTE Energy
Electric Distribution Operations

06/2013-08/2016
Occupant Safety Prototype and Testing Intern
ZF Friedrichshafen AG
Occupant Safety Systems and Driver Restraints, North America

Education

2018 M.S. Business Analytics
Broad Graduate School of Management
Michigan State University

2017 B.A. WITH HONORS in English Literature
College of Arts and Letters
Michigan State University

Service to the Profession

Current
Board Member, MSU Broad Graduate School of Management M.S. Business Data Science & Analytics Advisory Board

Former
Organizer, Ford Motor Company AI Paper Review Series
Planning Committee, Ford Motor Company R Users Guild
Planning Committee, Ford Motor Company Python Users Guild
Intern Recruiting Committee, Ford Motor Company
Intern Mentor, DTE Energy

Publications & Talks

PRESENTATIONS

Sep. 2021 **Presenter**, "Dangerous A.I.: A Primer on Algorithmic Bias and Digital Ethics," Data Science Topics Seminar Series, Broad Graduate School of Management, Michigan State University, East Lansing, MI.

Mar. 2021 **Keynote Speaker**, "Is Data Science Still 'Sexy'? A Recent History and Outlook of Analytics in Industry," Data Analytics and Software Event, Krannert School of Management, Purdue University, West Lafeyette, IN.

Feb. 2021 **Presenter**, "Connected Vehice Data Architecture for Large-scale Analytics," The Data Mine, Purdue University, West Lafeyette, IN.

Jan. 2021 **Presenter**, "Leveraging Flask for Microservice Development and Deployment," Cummins, Inc., Columbus, IN.

Dec. 2020 **Presenter**, "Applications of NLP for Service-oriented Companies," The Data Mine, Purdue University, West Lafeyette, IN.

Aug. 2020 **Presenter**, "Automating and Deploying Fundamental NLP Tasks," Global Data Insights & Analytics, Ford Motor Company, Dearborn, MI.

Apr. 2020 **Presenter**, "Addressing Structural Ambiguity in Automated Tabular Data Extraction," AI Advancement Center, Ford Motor Company, Dearborn, MI.

Jan. 2020- Apr. 2020 **Organizer and Presenter**, "Automotive Industry and Automobile Major Components Lecture Series," AI Advancement Center, Ford Motor Company, Dearborn, MI.

Feb. 2020 **Presenter**, "English is the New SQL: Translating Natural Language Questions to SQL," AI Advancement Center, Ford Motor Company, Dearborn, MI.
Selected as 1 of 10 talks to give at 2020 Global Data Insights & Analytics Conference (> 1,000 attendees); cancelled due to COVID-19 pandemic.

Nov. 2019 **Presenter**, "Information Retrieval for Question & Answering," Data Mine, Purdue University, West Lafeyette, IN.

Sep. 2019 **Co-Organizer and Presnter**, "Creating Effective Recommendation Engines using R," Ford R Guild Knowledge Share, Ford Motor Company, Dearborn, MI.

Posters

Dec. 2019 "Hidden Markov Models for Autonomous Vehicle Map Matching," 2019 Ford Global Control Conference, Ford Motor Company, Dearborn, MI.

Jun. 2019 "Connected Blue Zones: Crowdsourced Maps for Assisted Driving," 2019 Global Data Insights & Analytics Conference, Ford Motor Company, Livonia, MI.

Panels

May 2020 **Panelist**, "AI/ML Experts and Research Panel," Ford Product Development and Advanced Engineering, Ford Motor Company, Dearborn, MI.

Aug. 2018 **Panelist**, "Business Analytics Summer Internship Roundtable," Department of Accounting and Information Systems, Broad Graduate School of Management, Michigan State University, East Lansing, MI.

Papers, Journal Articles, & Technical Reports

2021 **Gould, J.** A Framework for Auditing Data Center Energy Usage and Mitigating Environmental Footprint. *arXiv*, Computer Science | Distributed, Parallel, and Cluster Computing. https://arxiv.org/abs/2102.04446

Teaching

Fall 2021 **Purdue University**
HONR 490: Special Topics in Data Science: Foundations of Geospatial Analytics
Upperclassman and Graduate-level Seminar

Research

01/2021-Present

Purdue University Research Comouting

Framework for Understanding and Mitigating Data Center Environmental Footprint

In coördination with Research Computing at Purdue University, our team's goal is to build a framework for universities to perform an audit of data center emissions, water and electrical usage, and operations costs to unerstand and reduce the environmental implications of running high-performance computing jobs, while contributing to university-driven Sustainability goals and standards.

08/2020-05/2021

Purdue University & Ford Motor Company

WHERE Clause Support for Natural Language Interface to Database Algorithms

This research project is an alliance between Ford Motor Company and Purdue University's Data Mine. Current approaches to the Natural Language Interface to Database (NLIDB) problem (an extension of Natural Language to SQL for relational databases) ignores WHERE clause prediction, making them unusable in industry. We aim to close this gap by developing a WHERE clause predictor to implement in models on Yale's Spider (Semantic Parsing and Text-to-SQL Challenge) task.

08/2019-05/2021

Purdue University & Ford Motor Company

In-Vehicle Multi-Turn Open Domain Question-Answer Systems

This research project is an alliance between Ford Motor Company and Purdue University's Data Mine. We created an in-vehicle Question-Answer system utilizing the 2020MY Lincoln Aviator manual. Given a question about a user's vehicle, which can be answered from the manual, the system searches the manual and returns the most probable answer, via ALBERT for text-retrieval and BERT- based Natural Language to SQL (NL2SQL) model for answers stored in tables.

04/2019-12/2019

University of Michigan & Ford Motor Company

Smart Vehicles for a Smart World

The Ford initiative with the University of Michigan's Transportation Research Institute (UMTRI) utilizes high-fidelity data collection technology from vehicles and intersections around Ann Arbor to collect $>$ 2 TB of data every month to find use cases in which Ford can use the vehicle as a sensor in various mobility products, such as finding correlations between driving behaviors, quantifying driving behavior, and classifying drivers into aggression segments for Usage-Based Insurance.

08/2017-12/2017

Michigan State University College of Arts and Letters

Translating and Analyzing Violent Language in a Vogelweide Poetry

The purpose of this senior research thesis, advised by Assistant Professor of Medieval Literature Tamar Boyadjian, was to translate Vogelweide's lyric poem *Ir Sult Sprechen Willekomen* from its original Middle High German to Modern German and analyze the implications of violent language used in themes of German Nationalism.

01/2017-05/2017

Michigan State University Digital Humanities Literary Cognition Lab

Neuroaesthetics: Poetry and Aesthetic Pleasure

This interdisciplinary project, led by MSU Professor of English Natalie Phillips, fo-

cuses on understanding the impact poetry has on a reader's brain, utilizing fMRI imaging to discover and quantify relationships.

Skills, Tools, etc.

Languages

English (native), German (advanced)

Programming

Python, R, Scala, SQL, CSS, HTML, Java, JavaScript, SAS, LaTeX

Tools/Software

Big Data: Hadoop, Spark, PySpark
Data Visualization: Tableau, PowerBI, Qlik
ML Framework: TensorFlow, PyTorch, Keras, Azure Cognitive Services
Geospatial: ArcGIS, Kinetica, GeoSpark, PostGIS, OSM
Data Management: Maximo, Alteryx, Postgres, MongoDB, SQLAlchemy, Azure
Data Workbench: Domino Data Labs, RAPIDS, Databricks
Orchestration and Deployment: Kubernetes, Django, Flask