mMLDS

MCCORMICK minor in
Machine Learning & Data Science

Northwestern | McCORMICK SCHOOL OF ENGINEERING

Agenda

- Introductions
- Definitions
- Program Overview
- Curriculum Requirements
- Applications & Admissions
- Q&A

Definitions

Data Science – A collection of tools and techniques that enable us to store and maintain data and derive meaning (or answer questions) from data.
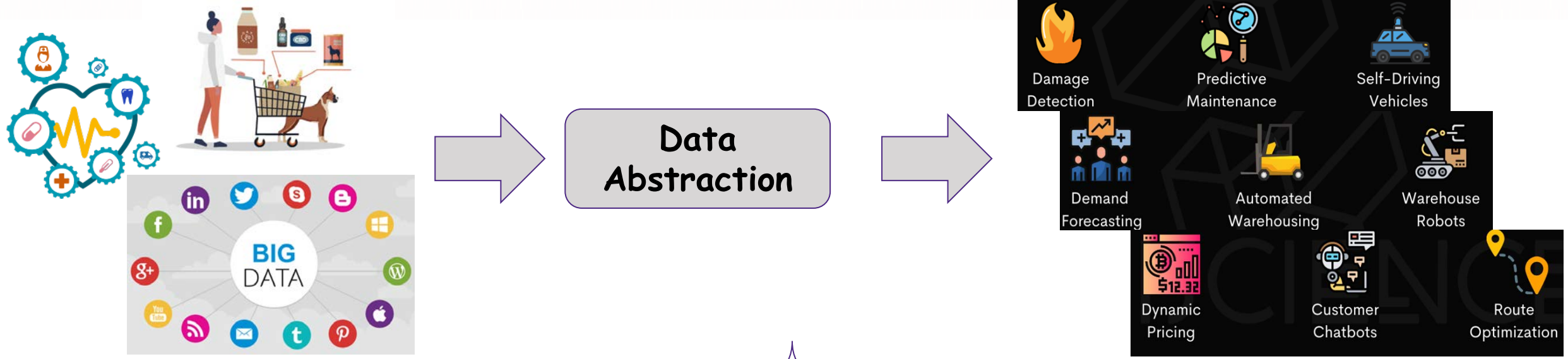
Artificial Intelligence – The study of building machines that can think or act humanly or rationally. This is commonly viewed as building machines that can do things that humans can do.
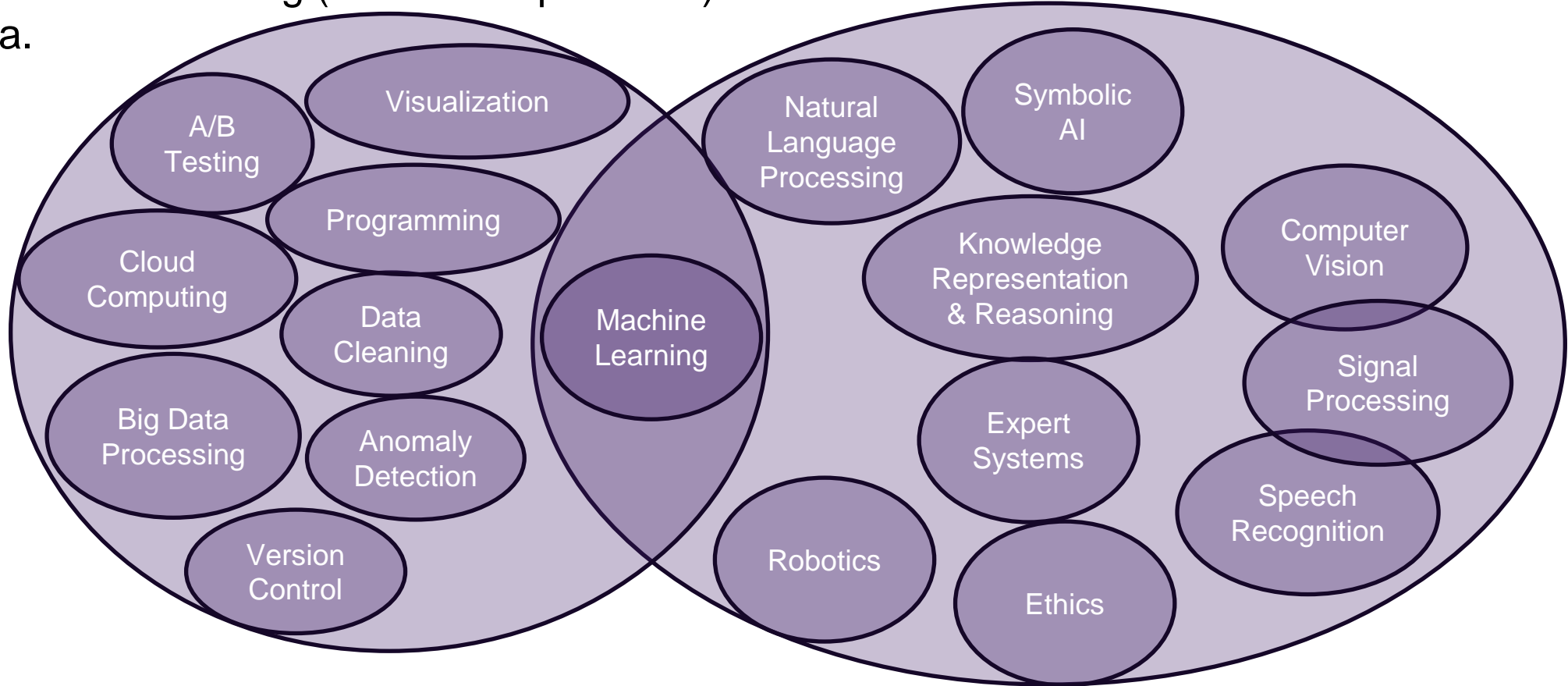
Machine Learning – A subfield of AI (& a tool for data scientists), ML is about designing algorithms to learn from data, that generalize to unseen data. Includes supervised, unsupervised & reinforcement learning.

Program Overview

- Practical knowledge fundamental to the data science lifecycle.

- Hands-on experience with models and techniques used for collecting, cleaning, and analyzing data

- Learn how to glean insights from data and think critically about data-driven decision making.

- Designed for undergraduate students across McCormick.

- Emphasize statistical and computational skills to glean and act on insights from data.

Data Abstraction

Damage Detection · Predictive Maintenance · Self-Driving Vehicles · Demand Forecasting · Automated Warehousing · Warehouse Robots · Dynamic Pricing · Customer Chatbots · Route Optimization

Cloud computing

Programming

Visualization

Machine learning

A/B testing

Data collection

Version control

Data cleaning

Big data processing

Anomaly detection

Northwestern | ENGINEERING
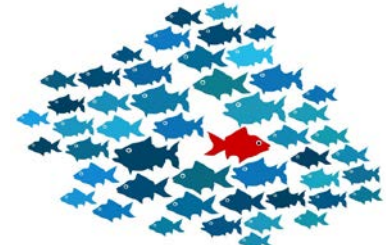
DATA_ENG 200: Foundations of Data Science

This course will introduce the steps of the data science lifecycle and common tools and techniques for data science. Here are some covered topics from the previous offering:

- Survival skills
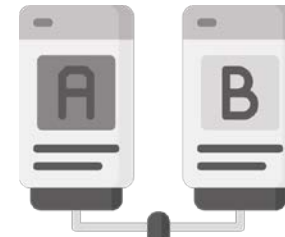- Preprocessing
- Basic analytics

Cloud computing   Programming   Visualization   Machine learning   A/B testing
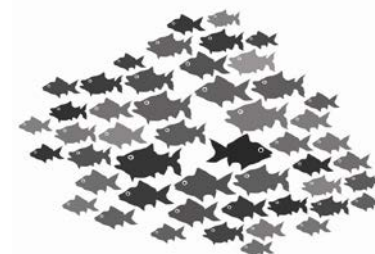
Data collection   Version control   Data cleaning   Big data processing   Anomaly detection

Northwestern | ENGINEERING

DATA_ENG 300: Data Engineering Studio

This course will teach students how to build a sustainable data science lifecycle. Here are some of the tentative topics:

- Machine learning
- Large scale processing
- Whole pipeline integration

Cloud computing   Programming   Visualization   Machine learning   A/B testing
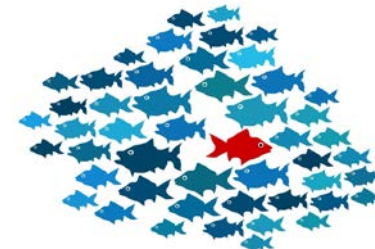
Data collection   Version control   Data cleaning   Big data processing   Anomaly detection

Northwestern | ENGINEERING

Data Science Intern

Oracle ★★★☆☆ 6,568 reviews

United States

Full-time, Internship

You must create an Indeed account before continuing to the company website to apply

Apply on company site ♡

What You'll Bring (Objective Minimum Qualifications):

To be considered for a Data Scientist Intern position, you must possess the Objective Minimum Qualifications (OMQs) below. **Please ensure that your application clearly indicates that you meet these OMQs** , either on your resume and/or by uploading additional documents such as a transcript.

- Will be enrolled in a Bachelor's, Master's or PhD degree program in an applied business or quantitative field such as Statistics, Operations Research, Applied Mathematics, Computer Science, Economics, Business Analytics, or equivalent field during the 2023-2024 school year.
- Have completed **at least your junior year** toward your undergraduate degree, or higher, by summer 2023.
- Have no more than **12 months** of professional full-time work experience in the technology field (excluding internships, research and/or teaching assistant roles, and military experience).
- Have academic coursework, projects, internships, and/or research experience with tools needed to execute the more technical aspects of statistical data analysis, including proficiency (e.g., can complete projects without any assistance) in **at least one** of the following programming languages:
  - Python, Java, R, SQL, PL/SQL.
- Have experience through coursework, projects, internships, and/or research with **at least one** of the following libraries:
  - Numpy, Scipy, Matplotlib, SciKit, TensorFlow, PyTorch, Spark
- Have completed coursework, projects, internships, and/or research in **two or more** of the following:
  - Algorithms
  - Artificial Intelligence
  - Big Data Query
  - Data Mining
  - Data Modeling
  - Machine Learning
  - Statistical Inference/Analysis

Data Scientist/ML Engineer

McDonald's Corporation ★★★☆☆ 213,693 reviews

Chicago, IL 60607

Full-time

Apply now ♡

Job Company

Qualifications

- BS in Computer Science, Mathematics, Economics, Statistics or a related discipline
- Strong experience in object-oriented programming using Python
- Experience developing, training, and evaluating (supervised/unsupervised) machine learning models such as random forest, SVMs, Naive Bayes, gradient boosting and kNN
- Experience developing, training, and evaluating deep-learning models using deep learning frameworks such as TensorFlow/Keras or Pytorch
- Project experience in computer vision such as object detection, image classification, text localization, OCR and semantic segmentation
- Experience with Sagemaker, S3, GCP Compute Engine, CloudGPUs, Cloud Storage, OpenCV, Boto3
- Experience building models which have been deployed to production, edge deployment preferred
- Strong ability to solve problems and to structure and simplify complex tasks
- Willingness to continue personal development in to stay at the cutting edge of data science
- Ability to build a sense of trust and rapport that creates a comfortable and effective workplace

Core Requirements: Four Courses

- **Programming Foundations:** COMP_SCI 150

- **Statistics Foundations:** Choose one from
  - BMD_ENG 220
  - CHEM_ENG 312
  - CIV_ENV 306
  - IEMS 201
  - IEMS 303

2023 DSE Requirements

- **MLDS Electives:** Choose two from an approved list
  - Access list with QR code
  - Some small changes to electives list under new requirements (remove CS 348; add CS 332 and other special topics).

Northwestern | ENGINEERING

- BMD_ENG 311 Computational Genomics
- BMD_ENG 312 Biomedical Applications in ML
- BMD_ENG 313 Wearable Devices
- CHEM_ENG 379-0 Computational Biology
- CIV_ENV 304-0 Systems Analysis
- CIV_ENV 377-0 Choice Modeling in Engineering
- CIV_ENV 480-1 Travel Demand Analysis & Forecasting 1
- CIV_ENV 495-0 Data Analytics for Transportation and Urban Infrastructure Applications
- ELEC_ENG 328-0 Information Theory & Learning
- ELEC_ENG 335 Deep Learning Foundations
- ELEC_ENG 373-0 Deep Reinforcement Learning
- ELEC_ENG 395-0 Optimization Techniques for ML
- ELEC_ENG 424-0 Distributed Optimization

- ELEC_ENG 433-0 Statistical Pattern Recognition
- ES_APPM 345-0 Applied Linear Algebra
- ES_APPM 375-1 Quantitative Biology I: Experiments, Data, Models, and Analysis
- ES_APPM 375-2 Quantitative Biology II: Experiments, Data, Models, and Analysis
- ES_APPM 472-0 Analysis of RNA Sequencing Data
- ES_APPM 479-0 Data Driven Methods for Dynamical Systems
- IEMS 307-0 Quality Improvement by Experimental Design
- IEMS 308-0 Data Science and Analytics
- IEMS 313-0 Foundations of Optimization
- IEMS 340-0 Qualitative Methods in Engineering Systems
- IEMS 341-0 Social Networks Analysis
- IEMS 351-0 Optimization Methods in Data Science

- MAT_SCI 358: Modeling and Simulation in Mat Sci
- MAT_SCI 391-0 Process Design
- MECH_ENG 301-0 Introduction to Robotics Laboratory
- MECH_ENG 329-0 Mechanistic Data Science
- MECH_ENG 341-0 Computational Methods for Engineering Design
- MECH_ENG 441-0 Engineering Optimization for Product Design and Manufacturing
- MECH_ENG 469-0 Machine Learning and Artificial Intelligence for Robotics
- MECH_ENG 495-0 Sensory Navigation and Machine Learning for Robotics

- COMP_SCI
  Agile Software, Deep Learning, Interactive Infor Viz, Computing Ethics, Viz for Scientific Comm, Rapid Prototyping, Online Markets, Causal Inference, Statistical Language Modeling, Natural & Artificial Vision

Specialization: Four Courses

Data Engineering	Hybrid	Machine Learning
COMP_SCI 217	COMP_SCI 214	COMP_SCI 111
IEMS 304	COMP_SCI 349	COMP_SCI 214
DATA_ENG 200	DATA_ENG 200	COMP_SCI 348
DATA_ENG 300	DATA_ENG 300	COMP_SCI 349
		not open to CS majors

- Specialization chosen at application: ML or DE/Hybrid

- DE/Hybrid limited by capacity of DATA_ENG courses; ML unconstrained.

- Waitlisted DE/Hybrid applicants will be offered ML option.

Northwestern | ENGINEERING

Data Studio Courses

- **DATA_ENG 200: Foundations of Data Science**
  Prerequisite: CS150, python
    - Winter 2025: TTh 9:30-10:50
    - Spring 2025: TTh 12:30-1:50

- **DATA_ENG 300: Data Engineering Studio**
  Prerequisites: Statistics Foundations course; CS150; CS 214 or 217; IEMS 304 or CS 349; DATA_ENG 200.
    - Winter 2025: TTh 12:30-1:50
    - Spring 2025: TTh 9:30-10:50

Important:

- Registration in DATA_ENG 200 is required in Winter or Spring 2025. Otherwise, you are welcome to apply in Fall 2025.

- Registration in DATA_ENG 300 will only be available in Spring 2025 if space is available after students in previous cohorts have registered. We will admit seniors for this year's cohort subject to available space.

Limitations on Double-Counting

- The minor must include **four** courses that are not used towards requirements for your major or any other programs.

  - "Major" means the 21-unit major requirements

  - Courses used for other McCormick requirements—unrestricted electives, math, basic science, theme, etc. are not considered to double count.

- Major-specific guidance for the current minor can be found on the DSE website. This applies to the DE/Hybrid tracks, but not yet to ML.

  - If you don't see yours, ask your advisor!
  - Updates for ML will be added when available.

To apply for Delta Cohort you must:

- Be a current McCormick undergraduate.

- Complete COMP_SCI 150 no later than Winter 2025.
  - Requires prerequisite of EA 1, CS 110, or CS 111

- Commit to taking DATA_ENG 200 in Winter or Spring 2025.
  - Depends on specialization—more on this in a moment.
  - Requires CS 150 as prerequisite.
  - Otherwise, you can apply in Fall 2025.

- Complete CS 150 before enrolling in Data_Eng 200.

- Have an expected graduation date no sooner than Winter 2026, when you will take DATA_ENG 300.
  - Graduating in 2025? More on this in a moment.

Students from ALL McCormick majors are welcome and encouraged to apply!

Northwestern | ENGINEERING

Applications and Admissions

- **Today:** Applications open
- **Sept 2:** Applications close
- **Sept 9:** First-round admissions decisions released
- **Sept 15:** Deadline to confirm participation
- **Sept 18:** Waitlist admissions decisions released
- **Sept 24:** Fall 2024 classes begin; deadline for waitlisted students to confirm.

Questions?

Still have Questions?

Email dse@northwestern.edu
Or
dseadmin@u.northwestern.edu