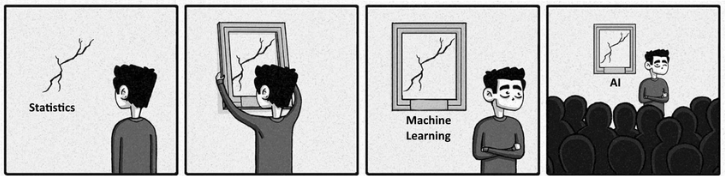
Machine Learning

Data Analysis for Journalism and Political Communication
(Spring 2026)

Prof. Bell



Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.

Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.
- **Statistics** primarily focuses on **inference**.

Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.
- **Statistics** primarily focuses on **inference**.
  - Goal: To use a sample to understand and draw conclusions about the broader population it came from.

Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.
- **Statistics** primarily focuses on **inference**.
  - Goal: To use a sample to understand and draw conclusions about the broader population it came from.
- **Machine Learning** primarily focuses on **prediction**.

Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.
- **Statistics** primarily focuses on **inference**.
  - Goal: To use a sample to understand and draw conclusions about the broader population it came from.
- **Machine Learning** primarily focuses on **prediction**.
  - Goal: To build a model on a sample of "training" data that makes the most accurate predictions possible on new, unseen "test" data.

Machine Learning vs. Statistics

- Both fields use data and algorithms to find patterns. The main difference lies in their primary goal: **inference** or **prediction**.
- **Statistics** primarily focuses on **inference**.
  - Goal: To use a sample to understand and draw conclusions about the broader population it came from.
- **Machine Learning** primarily focuses on **prediction**.
  - Goal: To build a model on a sample of "training" data that makes the most accurate predictions possible on new, unseen "test" data.
- Despite different goals, core principles like avoiding selection bias, ensuring data quality, and understanding sample limitations are critical in both fields.

A Key ML Practice: Training and Testing

- In ML, we split our data into a **training set** to build the model and a separate **test set** to evaluate its predictive performance.

A Key ML Practice: Training and Testing

- In ML, we split our data into a **training set** to build the model and a separate **test set** to evaluate its predictive performance.
- A flexible model can learn the deep, underlying patterns (**signal**) in the training data, but it can also learn the random, irrelevant quirks (**noise**).

A Key ML Practice: Training and Testing

- In ML, we split our data into a **training set** to build the model and a separate **test set** to evaluate its predictive performance.
- A flexible model can learn the deep, underlying patterns (**signal**) in the training data, but it can also learn the random, irrelevant quirks (**noise**).
- **Overfitting** occurs when the model learns the noise instead of just the signal. It memorizes the training data too closely.

A Key ML Practice: Training and Testing

- In ML, we split our data into a **training set** to build the model and a separate **test set** to evaluate its predictive performance.
- A flexible model can learn the deep, underlying patterns (**signal**) in the training data, but it can also learn the random, irrelevant quirks (**noise**).
- **Overfitting** occurs when the model learns the noise instead of just the signal. It memorizes the training data too closely.
  - An overfit model performs great on the training data but fails to make accurate predictions on the new test data. It doesn't **generalize** well.

A Key ML Practice: Training and Testing

- In ML, we split our data into a **training set** to build the model and a separate **test set** to evaluate its predictive performance.
- A flexible model can learn the deep, underlying patterns (**signal**) in the training data, but it can also learn the random, irrelevant quirks (**noise**).
- **Overfitting** occurs when the model learns the noise instead of just the signal. It memorizes the training data too closely.
  - An overfit model performs great on the training data but fails to make accurate predictions on the new test data. It doesn't **generalize** well.
- **Underfitting** is the opposite problem: the model is too simple and fails to capture the underlying signal in the data.

Machine Learning and AI

- Machine learning models learn patterns directly from data, rather than being explicitly programmed with rules. This process is often categorized into three main types:

Machine Learning and AI

- Machine learning models learn patterns directly from data, rather than being explicitly programmed with rules. This process is often categorized into three main types:
  - **Supervised Learning**: The model learns from **labeled** data (data with known outcomes) to predict outcomes for new data.
  - **Unsupervised Learning**: The model works with **unlabeled** data to discover hidden patterns or structures on its own.
  - **Reinforcement Learning**: A model (or "agent") learns by interacting with an environment. It performs actions and receives rewards or penalties, learning the best strategy over time (e.g., an AI learning to play chess).

Machine Learning and AI

- Machine learning models learn patterns directly from data, rather than being explicitly programmed with rules. This process is often categorized into three main types:
    - **Supervised Learning**: The model learns from **labeled** data (data with known outcomes) to predict outcomes for new data.
    - **Unsupervised Learning**: The model works with **unlabeled** data to discover hidden patterns or structures on its own.
    - **Reinforcement Learning**: A model (or "agent") learns by interacting with an environment. It performs actions and receives rewards or penalties, learning the best strategy over time (e.g., an AI learning to play chess).
- As models become more complex, it can be difficult to understand *why* they make a specific prediction. This challenge is known as **explainability** or the "black box" problem.

Bias in ML Models

- ML models learn from the data they are given. If that data reflects historical human biases, the model will learn and often **amplify** those biases.

Bias in ML Models

- ML models learn from the data they are given. If that data reflects historical human biases, the model will learn and often **amplify** those biases.
- This is a supercharged version of the "Garbage In, Garbage Out" principle. The model treats biased historical patterns as the truth to be optimized.

Bias in ML Models

- ML models learn from the data they are given. If that data reflects historical human biases, the model will learn and often **amplify** those biases.
- This is a supercharged version of the "Garbage In, Garbage Out" principle. The model treats biased historical patterns as the truth to be optimized.
- **Public Policy Example: Criminal Justice**: The COMPAS algorithm, used to predict the likelihood of a defendant re-offending, was found to be biased against Black defendants. It was more likely to incorrectly flag them as high-risk compared to white defendants.

Bias in ML Models

- ML models learn from the data they are given. If that data reflects historical human biases, the model will learn and often **amplify** those biases.
- This is a supercharged version of the "Garbage In, Garbage Out" principle. The model treats biased historical patterns as the truth to be optimized.
- **Public Policy Example: Criminal Justice**: The COMPAS algorithm, used to predict the likelihood of a defendant re-offending, was found to be biased against Black defendants. It was more likely to incorrectly flag them as high-risk compared to white defendants.
- **Other Examples**: AI-powered hiring tools that discriminate against female candidates because they were trained on historical hiring data from a male-dominated industry, or mortgage approval algorithms that perpetuate racial redlining.

Large Language Models (LLMs)

- LLMs are a type of deep learning model trained on vast amounts of text data. At their core, they are sophisticated pattern-matching systems.

Large Language Models (LLMs)

- LLMs are a type of deep learning model trained on vast amounts of text data. At their core, they are sophisticated pattern-matching systems.
- They generate output by probabilistically predicting the next “token” (a word or part of a word) in a sequence, given the text that came before it.

Large Language Models (LLMs)

- LLMs are a type of deep learning model trained on vast amounts of text data. At their core, they are sophisticated pattern-matching systems.
- They generate output by probabilistically predicting the next "token" (a word or part of a word) in a sequence, given the text that came before it.
- A key limitation is that an LLM's knowledge is frozen in time; it ends when its training data ends. It has no inherent knowledge of events that occur after its training is complete.

Large Language Models (LLMs)

- LLMs are a type of deep learning model trained on vast amounts of text data. At their core, they are sophisticated pattern-matching systems.
- They generate output by probabilistically predicting the next "token" (a word or part of a word) in a sequence, given the text that came before it.
- A key limitation is that an LLM's knowledge is frozen in time; it ends when its training data ends. It has no inherent knowledge of events that occur after its training is complete.
- To overcome this, LLMs can be used as part of an **agent** system. These agents can use **tool calls** to access external, real-time information (like a search engine) or perform actions, allowing them to provide up-to-date responses.

Key Challenges and Risks with LLMs

- **Hallucinations**: LLMs can confidently invent facts, sources, and quotes. This is not a bug, but a natural result of how they work.

Key Challenges and Risks with LLMs

- **Hallucinations**: LLMs can confidently invent facts, sources, and quotes. This is not a bug, but a natural result of how they work.
- **Bias at Scale**: LLMs are trained on vast sections of the internet, which contains immense amounts of bias. They absorb and can reproduce racism, sexism, and other biases.

Key Challenges and Risks with LLMs

- **Hallucinations**: LLMs can confidently invent facts, sources, and quotes. This is not a bug, but a natural result of how they work.
- **Bias at Scale**: LLMs are trained on vast sections of the internet, which contains immense amounts of bias. They absorb and can reproduce racism, sexism, and other biases.
- **Misinformation**: Because they are so good at generating plausible text, LLMs are powerful tools for creating and spreading disinformation and propaganda cheaply and at massive scale.

Key Challenges and Risks with LLMs

- **Hallucinations**: LLMs can confidently invent facts, sources, and quotes. This is not a bug, but a natural result of how they work.
- **Bias at Scale**: LLMs are trained on vast sections of the internet, which contains immense amounts of bias. They absorb and can reproduce racism, sexism, and other biases.
- **Misinformation**: Because they are so good at generating plausible text, LLMs are powerful tools for creating and spreading disinformation and propaganda cheaply and at massive scale.
- **Data Provenance**: The data used to train these models is often scraped from the web without permission, raising major copyright and ethical questions that are still being debated in court.

LLMs and Copyright: Key Lawsuits

- The training of LLMs on vast datasets of internet content, including copyrighted material, has led to significant legal challenges.

LLMs and Copyright: Key Lawsuits

- The training of LLMs on vast datasets of internet content, including copyrighted material, has led to significant legal challenges.
- **Authors Guild v. OpenAI (and others)**
  - **Who**: Prominent authors (e.g., George R.R. Martin, John Grisham) and the Authors Guild against OpenAI and other AI developers.
  - **Claim**: AI companies copied copyrighted books without permission to train LLMs, and these LLMs can generate content that competes with or is derivative of their work.

LLMs and Copyright: Key Lawsuits

- The training of LLMs on vast datasets of internet content, including copyrighted material, has led to significant legal challenges.
- **Authors Guild v. OpenAI (and others)**
  - **Who**: Prominent authors (e.g., George R.R. Martin, John Grisham) and the Authors Guild against OpenAI and other AI developers.
  - **Claim**: AI companies copied copyrighted books without permission to train LLMs, and these LLMs can generate content that competes with or is derivative of their work.
- **The New York Times Co. v. OpenAI and Microsoft**
  - **Who**: The New York Times Company against OpenAI and Microsoft.
  - **Claim**: Defendants used millions of NYT articles to train LLMs, which now produce output that closely mimics NYT content, sometimes reproducing it verbatim, and could draw readers away from the newspaper.

Build Your Own ML Model

Goal: Can we predict the tone of a campaign ad about a candidate based on how the candidate is portrayed in the ad?

Step 1: Gather Your Training Data

- Go to The Living Room Candidate (https://www.livingroomcandidate.org/)
- Select ads from any election(s) **except** 2024.
- Take 10-15 screen grabs that clearly feature a candidate.
- Create two folders: “Positive Tone” and “Negative Tone.” Place each screen grab in the appropriate folder based on whether the image portrays the candidate positively or negatively within the ad.

Build Your Own ML Model

Step 2: Train Your Model

- Go to Teachable Machine (https://teachablemachine.withgoogle.com)
- Go to Image Project → Standard Image Model.
- Upload your "Positive Tone" images to Class 1, and your "Negative Tone" images to Class 2.
- Click "Train Model."

Step 3: Test and Evaluate

- I will provide you with a test set of 10 images from 2024 campaign ads.
- Use Teachable Machine's file upload feature to test your model on each of these 10 images.
- How did your model do? Where did it do well, and where did it struggle? Why do you think this happened?