

Lessons from the Future

Predictions in Finance, Sports and Medicine

Dr. Robert P. Schumaker
Associate Professor of Computer Science
Director – Data Analytics Lab

Lessons from the Future

- What is a Data Scientist?
 - Solve large, complex, often unstructured or ill-defined problems using a data-driven approach
 - Able to predicting customer needs and anticipate orders
 - Someone that sees patterns, subtle details, looks at the world differently

Lessons from the Future

- Is Data Science just Machine Learning?
 - Machine Learning is a part of the process
 - 1. Must understand the problem
 - 2. Must know what ML algorithm to use based on its strengths and limitations
 - 3. MUST BE ABLE TO INTERPRET THE RESULTS

Lessons from the Future

- What is Data Science?
 - Knowledge of Statistics – Statistician
 - Clean and Analyze Data – Business Analyst
 - Able to Data Mine – Entry-level Data Scientist
 - Translate Business Requirements – Data Scientist
 - System Builder – Senior Data Scientist

Lessons from the Future

- The Data Science Approach
 - Identify a general area to study
 - Acquisition
 - Determine what data resources are available
 - Collect data directly or through API
 - Clean the relevant data
 - Representation
 - Ascertain the independent and dependent variables
 - Shape the research questions
 - Consider ways to model the data

Lessons from the Future

- The Data Science Approach
 - Description / Prediction
 - Determine the relevant statistics, machine learning algorithms
 - Train, Test and Validate
 - Refine research questions, seek additional data sources, iterate
 - Delivery
 - Prepare a technical research report to communicate results
 - Present the findings
 - Create a software prototype

Lessons from the Future

- Lessons in Finance (Textual/Financial Prediction)
 - How it started

THE WALL STREET JOURNAL.
© 1987 Dow Jones & Company, Inc. All Rights Reserved.

VOL. CCX NO. 79 *** EDITION: Eastern *** TUESDAY, OCTOBER 20, 1987 PRINCETON, New Jersey *** 50 CENTS

Risky Strategy

Attack on Iranian Site Wins Praise but Means Violence May Escalate

U.S. Says Strike Shows It Is In Persian Gulf to Stay; Congress Remains Uneasy

Soviet Role Is Called Unlikely

By Tim Condon and
Jeff Gerson, F. Ben
Joshi, Kenneth F. Smith, and
Washington Post Staff Writers

WASHINGTON—The American strike against an Iranian offshore military site, which won widespread support in Congress, where it was seen as a test of the president's resolve, also drew criticism from some quarters for its potential to escalate violence.

The attack, which was greeted warmly by many in the Persian Gulf, was seen as a test of the president's resolve, and it was widely expected that it would lead to a more aggressive U.S. posture in the region.

For some, the strike was a signal that the U.S. was serious about its commitment to the Persian Gulf. It was seen as a test of the president's resolve, and it was widely expected that it would lead to a more aggressive U.S. posture in the region.

Some, however, were concerned that the strike might lead to a more aggressive U.S. posture in the region. They argued that the strike might lead to a more aggressive U.S. posture in the region.

Some, however, were concerned that the strike might lead to a more aggressive U.S. posture in the region. They argued that the strike might lead to a more aggressive U.S. posture in the region.

What's News—

Business and Finance

THE STOCK MARKET CRASHED

as panic selling swept the Dow Jones Industrial Average down 508 points, or 22.8%, to 2154.74. The record decline far exceeded the drop of Oct. 29, 1929, when the average lost 23.8%. Most other market indicators also slumped to record lows, as Big Board volume soared to 644.3 million shares, well above the previous record.

Investors don't expect the market's decline to usher in a depression, mostly due to its departure from the banking and financial systems. But many small investors remain worried.

Investors don't expect the market's decline to usher in a depression, mostly due to its departure from the banking and financial systems. But many small investors remain worried.

World-Wide

THE U.S. BOMBARDERS

on October 18, the strike by Navy warships, ordered by Reagan in retaliation for Friday's attack on a U.S.-flagged Kuwait oil tanker by Tehran's forces, damaged the platform, which had been converted to a military command center in the Gulf. The U.S. had broken off warships prior to the raid. American forces also destroyed two Iranian radar and communication equipment, according to the Pentagon. The attack was greeted by the U.S. as a military victory, but some in the administration remain worried.

Investors don't expect the market's decline to usher in a depression, mostly due to its departure from the banking and financial systems. But many small investors remain worried.

Market Fallout

The hours run wild on Wall Street in the wake of the market's worst crash since 1929. The reaction by market participants and investors was one of disbelief. Investors include:

- Trading on the floor of the New York Stock Exchange was a chaotic scene as investors tried to get out of their positions. The trading floor was a scene of chaos as investors tried to get out of their positions.
- The American bond market was hit hard, with yields rising sharply.
- The American bond market was hit hard, with yields rising sharply.
- The American bond market was hit hard, with yields rising sharply.

The Market Debate

Routes Worst Fears Of Little Investors

Many Are Beginning to Talk In Terms of Doom; Big Investors Are Bigger

A Wall Street Journal reporter wrote: "The stock market's decline has been a disaster for many investors, but it has also been a disaster for many investors."

The stock market's decline has been a disaster for many investors, but it has also been a disaster for many investors.

The Big Crash

Percentage Decline Is Far Steeper Than '29; Bond Prices Surge

NEW YORK — The stock market crashed yesterday.

The Dow Jones Industrial Average plummeted 508 points, or 22.8%, to 2154.74. The drop far exceeded the decline of Oct. 29, 1929, when the average lost 23.8%. Most other market indicators also slumped to record lows, as Big Board volume soared to 644.3 million shares, well above the previous record.

Investors don't expect the market's decline to usher in a depression, mostly due to its departure from the banking and financial systems. But many small investors remain worried.

A Repeat of '29? Is Not Expected

Banking System Safeguards And Federal Mechanisms Are Viewed as Adequate

By Kenneth H. Rogoff
Staff Reporter of The Wall Street Journal

Can it happen again?

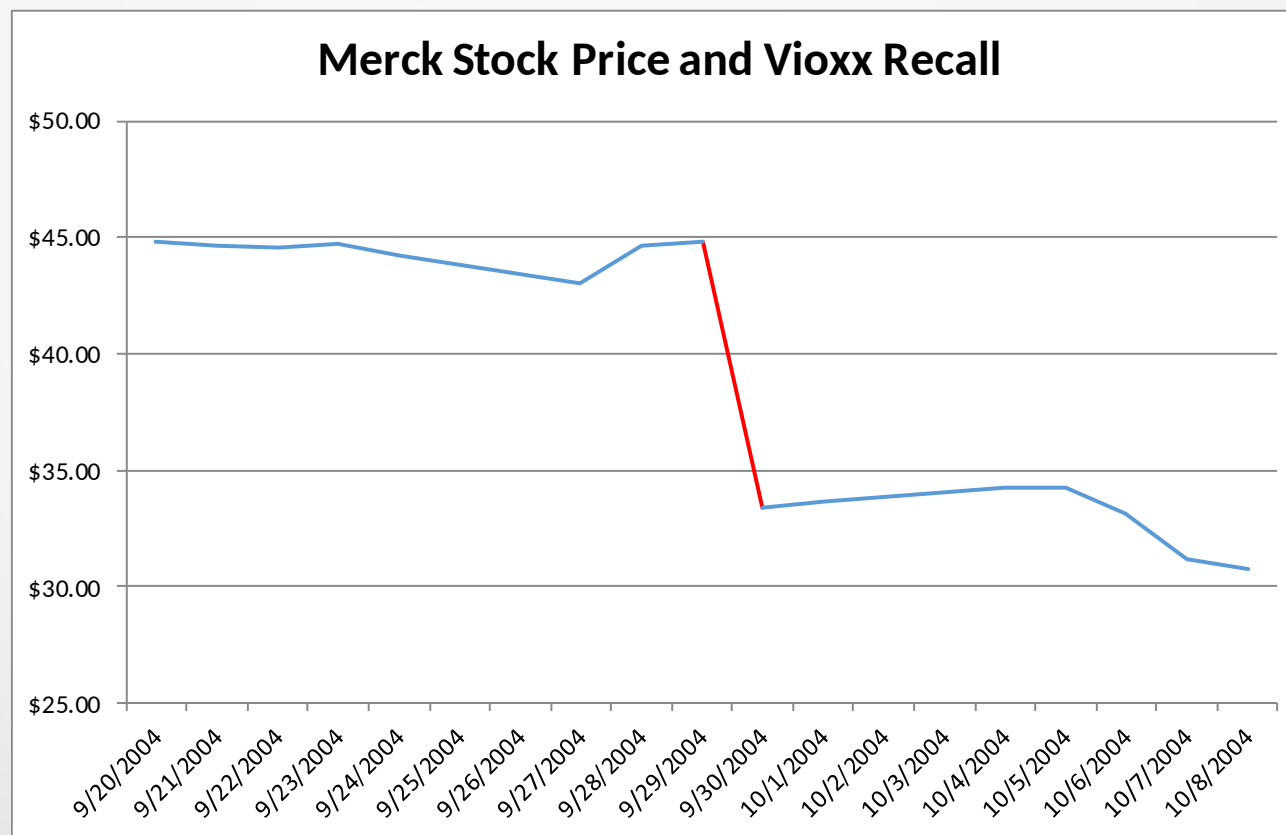
On Oct. 29, 1929, the stock market fell 23.8%, ushering in the Great Depression. While the market plunged 22.8% yesterday, many investors are not expecting a repeat of the 1929 crash.

Many investors are not expecting a repeat of the 1929 crash. They argue that the current market conditions are different from those of 1929.

Many investors are not expecting a repeat of the 1929 crash. They argue that the current market conditions are different from those of 1929.

Lessons from the Future

- Textual/Financial Prediction
 - Building a news-aware trader



Lessons from the Future

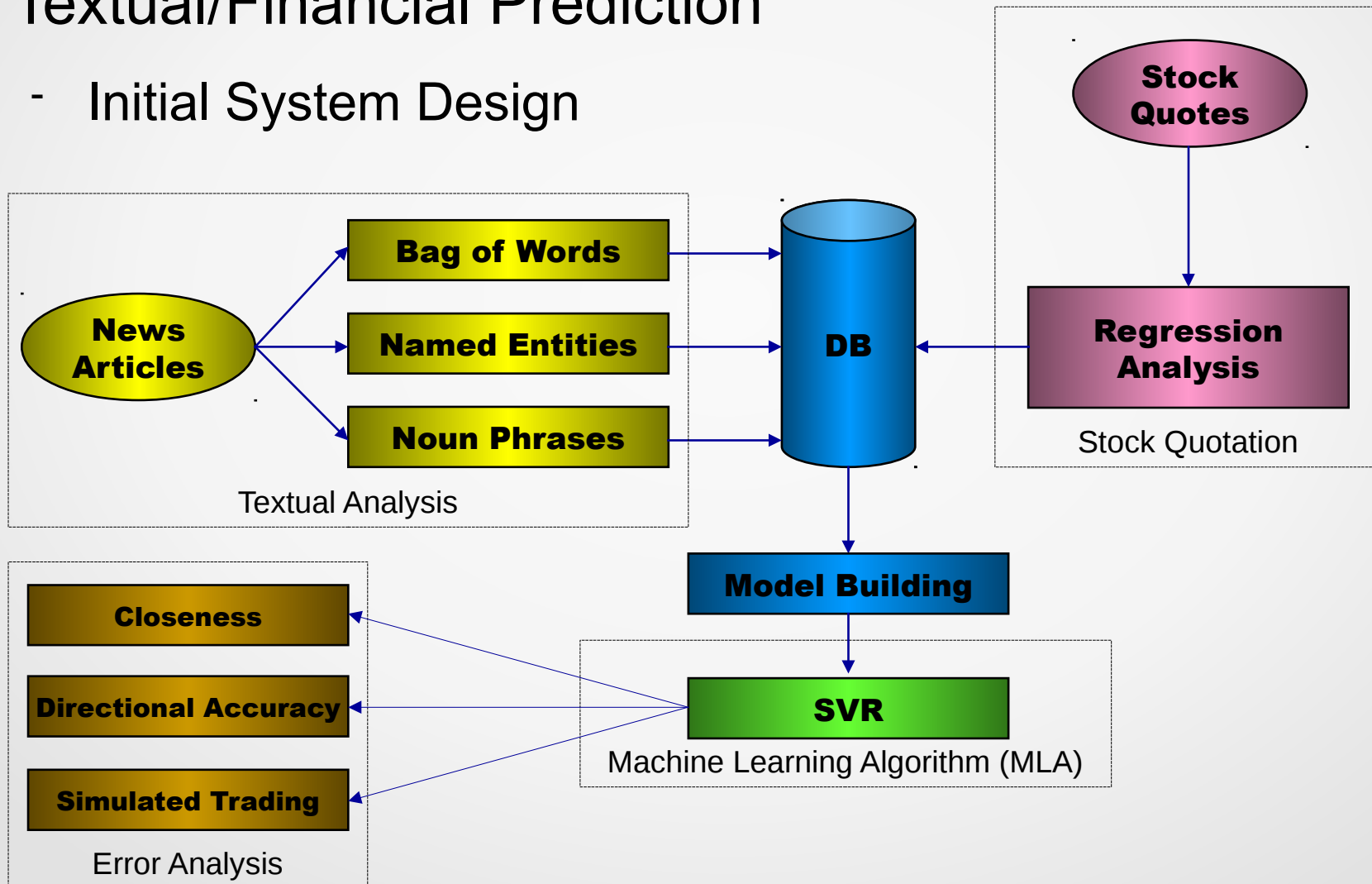
- Textual/Financial Prediction

- Problems

- Prior research focused on price direction prediction / classification
 - Prior research used only Bag of Words
 - Very limited work, lacked critical details
 - No research built a simulated trading engine (HFT)

Lessons from the Future

- Textual/Financial Prediction
 - Initial System Design



Lessons from the Future

- Textual/Financial Prediction
 - Later System Designs
 - Similarity in industry classification
 - Comparison to trading experts
 - Building a portfolio
 - Adding dimensions of sentiment
 - Stock Price Management
 - Should the article be released
 - When should it be released
 - What media should it be released in
 - *What terms should be used*

Lessons from the Future

- Textual/Financial Prediction
 - Lessons learned
 - Data Science takes time
 - Need to be cross-disciplinary
 - Data-driven leads to unexpected insights
 - Refining variables and understanding a entire solution landscape can take a lifetime

Lessons from the Future

- Lessons in Sports (Sport Prediction)
 - How it started



Lessons from the Future

- Sport Prediction

- Problems

- Prior research focused on binary classification
 - Prior research manually curated inputs
 - Very limited work, **but had the critical details**
 - Small tweaks to variables, algorithms and domains

Lessons from the Future

- Lessons in Sports (Wagering Prediction)
 - How it started



#nufc

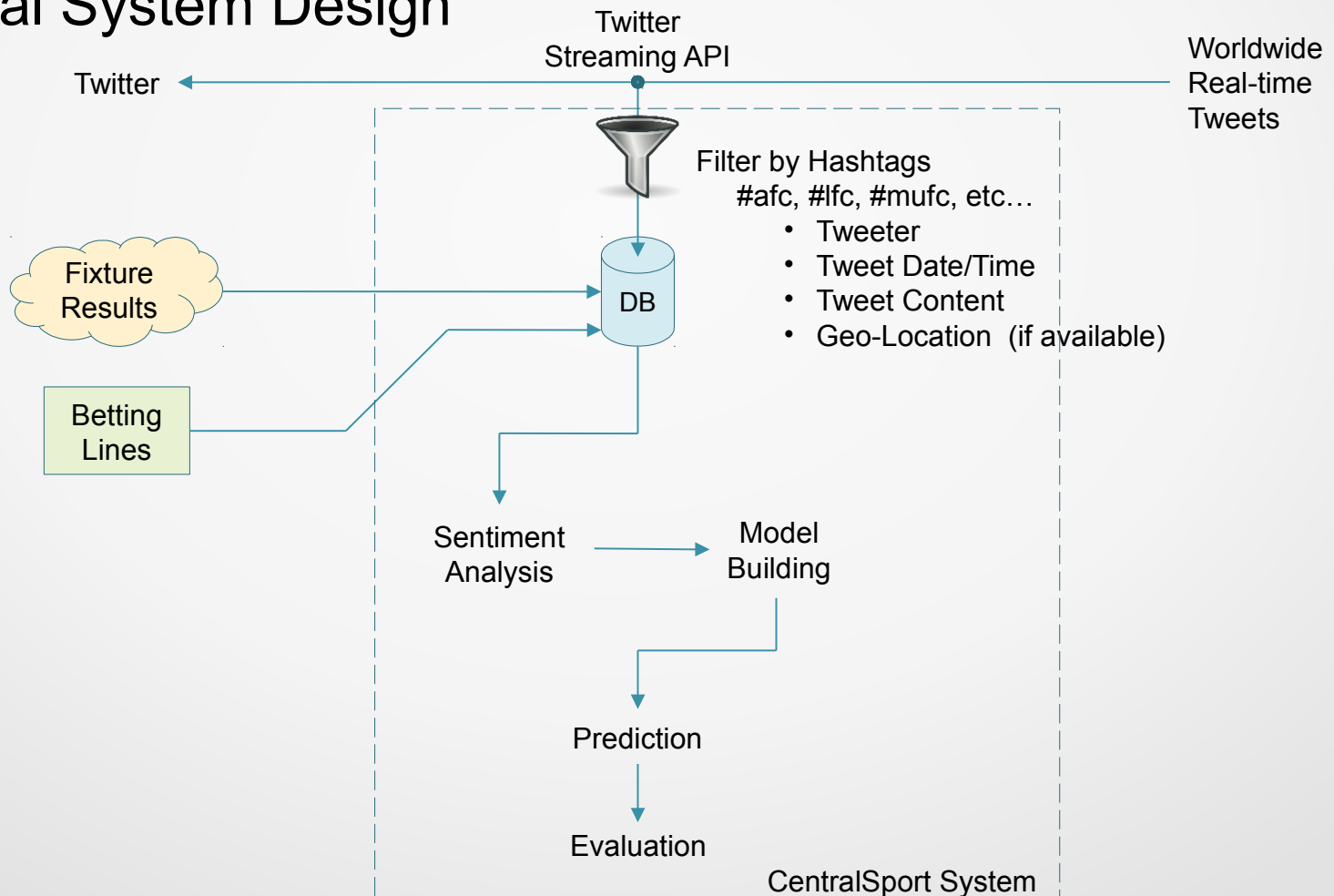
Supporting Newcastle is actually making me hate football.

Lessons from the Future

- Wagering Prediction
 - Problems
 - Prior research focused on binary classification
 - Prior research manually curated inputs
 - Very limited work, lacked critical details
 - No correlation to wagering

Lessons from the Future

- Wagering Prediction
 - Initial System Design



Lessons from the Future

- Sports and Wagering Prediction
 - Lessons learned
 - Need to be cross-disciplinary
 - Data-driven leads to unexpected insights
 - Refining variables and understanding a entire solution landscape can take a lifetime
 - Tweeters from Sunderland are not nice people

- How it started

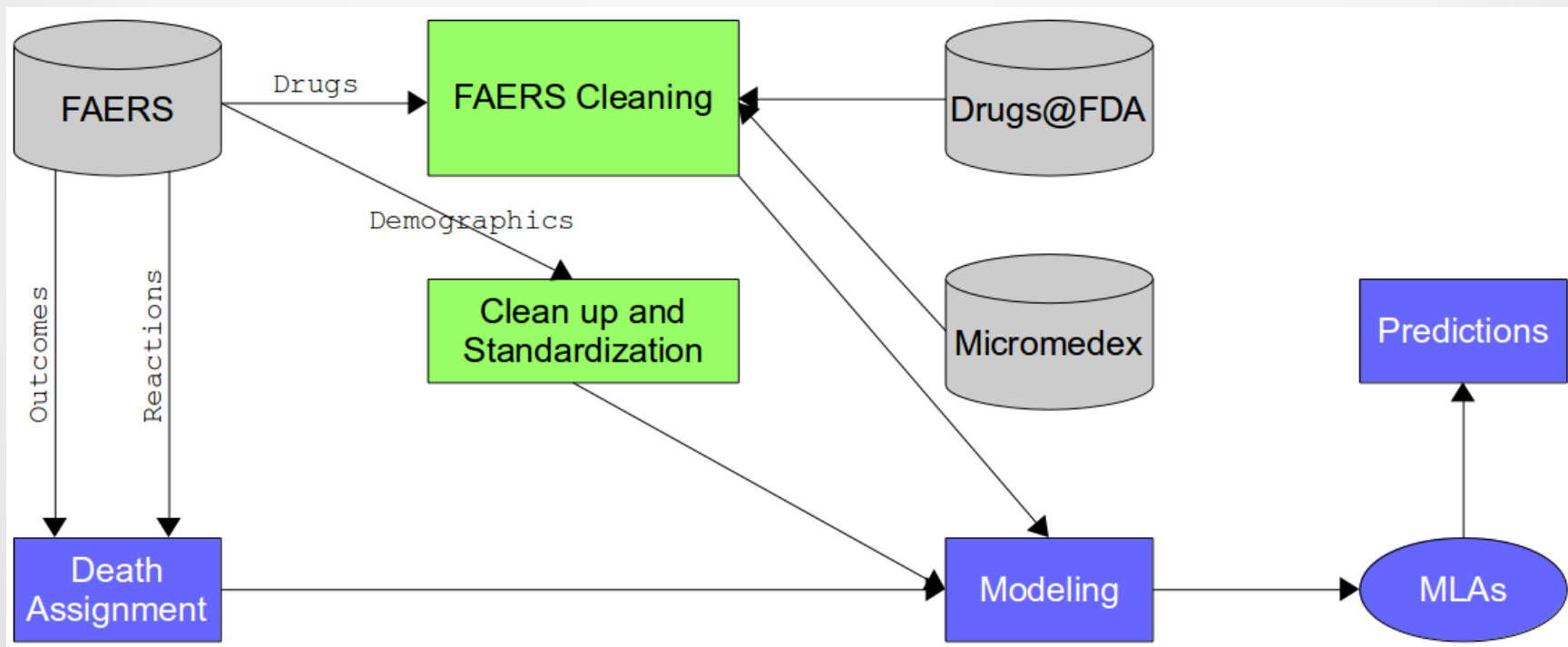


Lessons from the Future

- ADE Prediction
 - Problems
 - Prior research focused on 2 drug combinations
 - Available data is very large and dirty
 - No related work, lacked critical details

Lessons from the Future

- ADE Prediction
 - Initial System Design



Lessons from the Future

- ADE Prediction
 - Current System Designs
 - Look at reactions in combination with opioids
 - Partition based on patient demographics
 - Partition based on disease

Lessons from the Future

- ADE Prediction

- Lessons learned

- Data Science takes time
 - Need to be cross-disciplinary
 - Data-driven leads to unexpected insights
 - Refining variables and understanding a entire solution landscape can take a lifetime
 - Be flexible in system building
 - Don't take Heroin and Buprenorphine together

Lessons from the Future

- Lessons I've Learned on Becoming a Data Scientist
 - Learn as much as you can
 - Keep up-to-date
 - Don't be afraid of hard problems
 - There are no shortcuts, it will take time
 - Go where the data takes you
 - Make it fun