Georgia Institute of Technology

Dr. Tucker Balch
Associate Professor
School of Interactive Computing

Computational Investing, Part I

221: Intro to Machine Learning

Find out how modern electronic markets work, why stock prices change in the ways they do, and how computation can help our understanding of them. Learn to build algorithms and visualizations to inform investing practice.

School of Interactive Computing

The Market is a “Big Data” Challenge

- **13,000 securities** trading in America
- **120** technical & fundamental factors
- **252** trading days per year

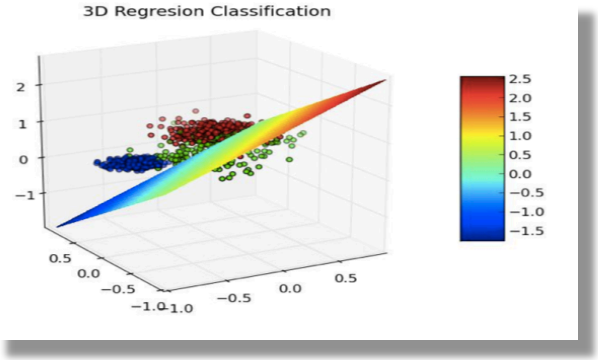
= 1.96 Billion data points in 5 years

Machine learning can help!

What is Machine Learning?

"The construction and study of systems that can learn from data."

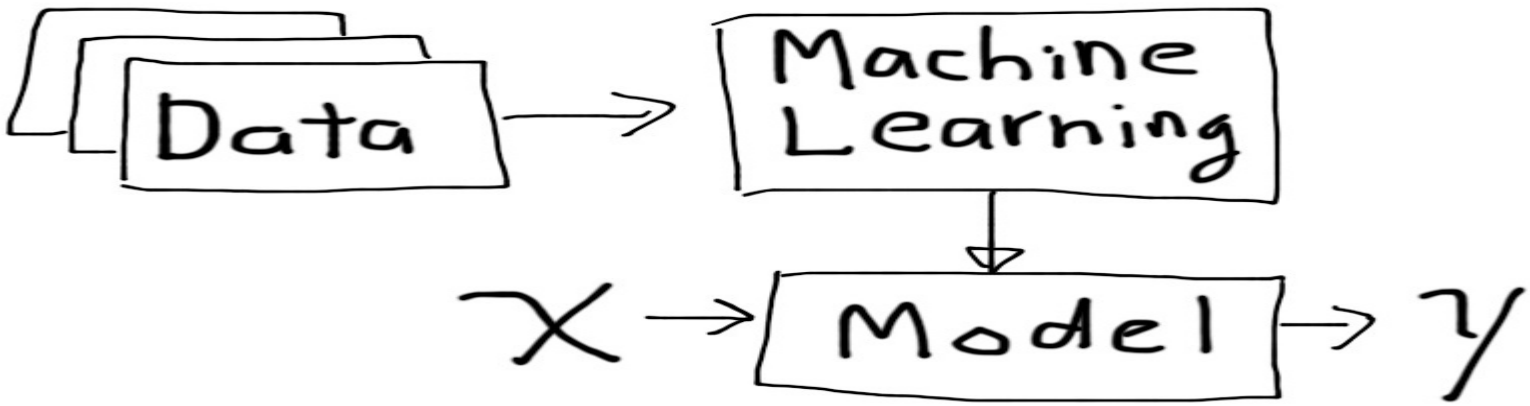
--Wikipedia



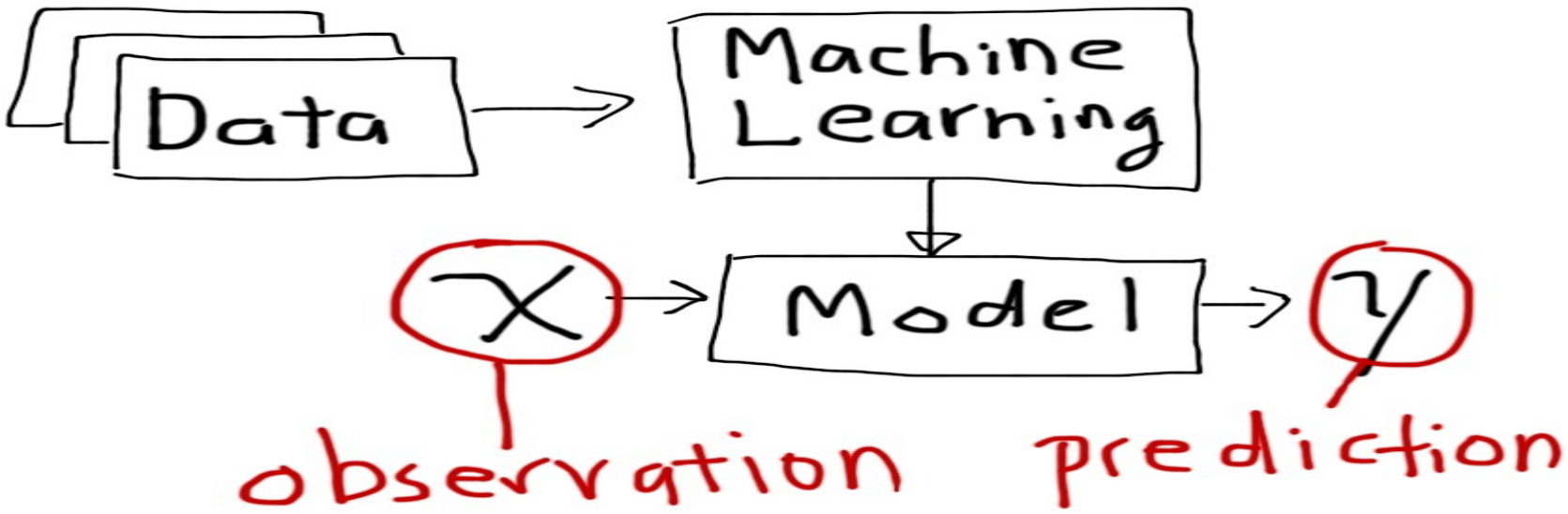
Major uses

- Classification:
  - What is it?
- Regression
  - A numerical prediction

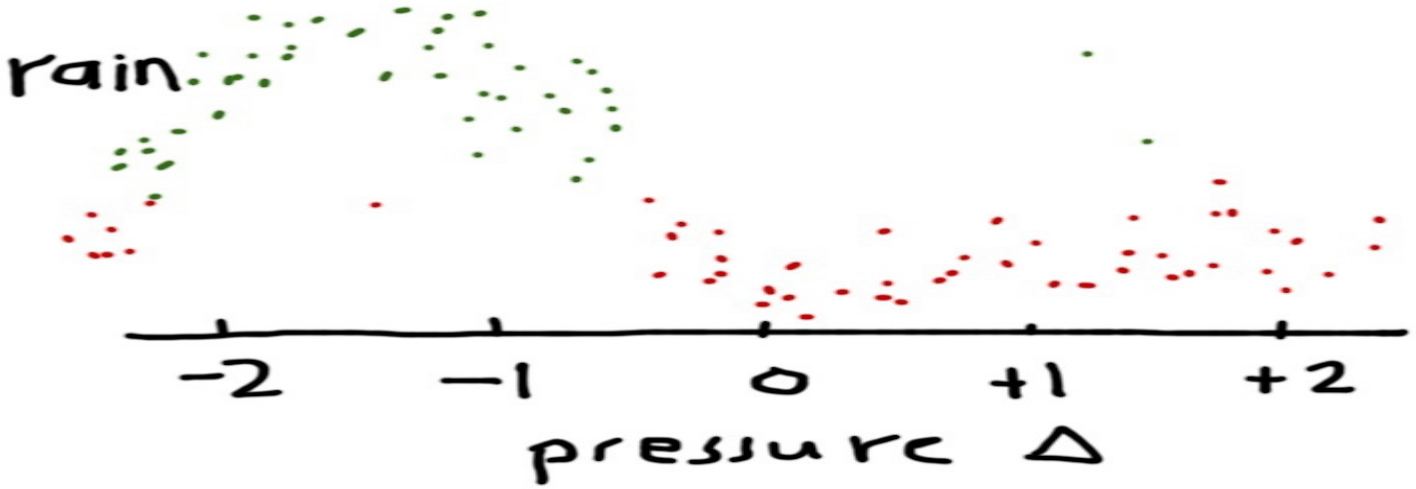
Building a Model From Data



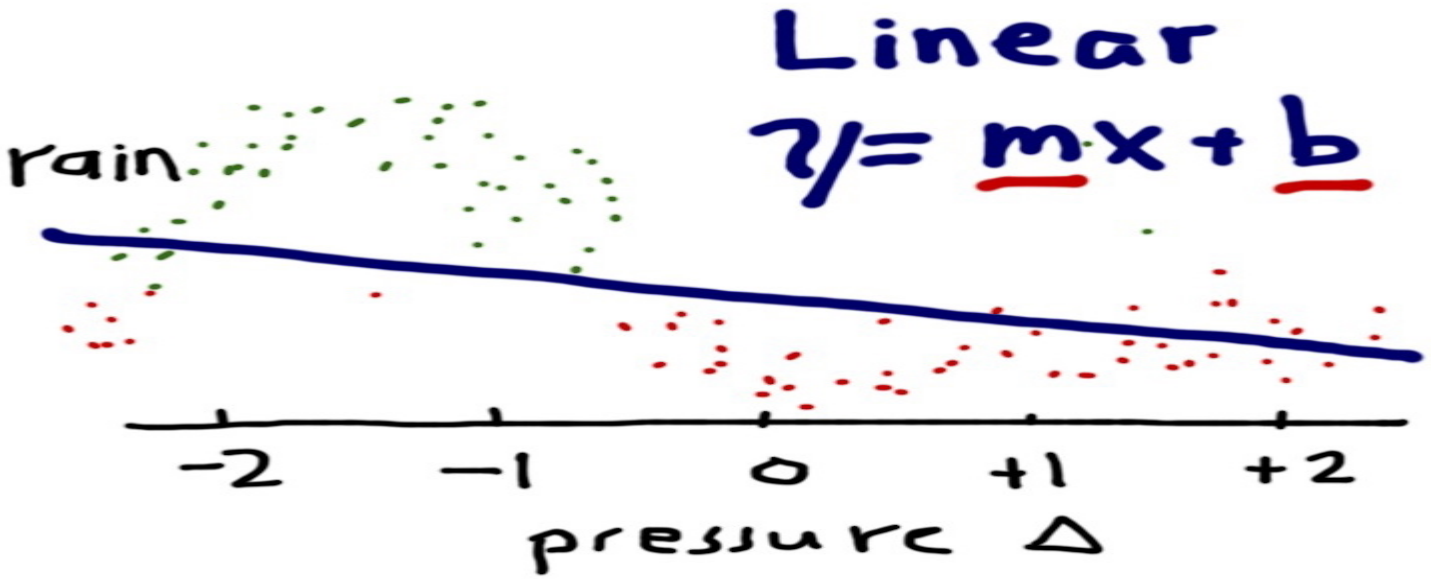
Building a Model From Data



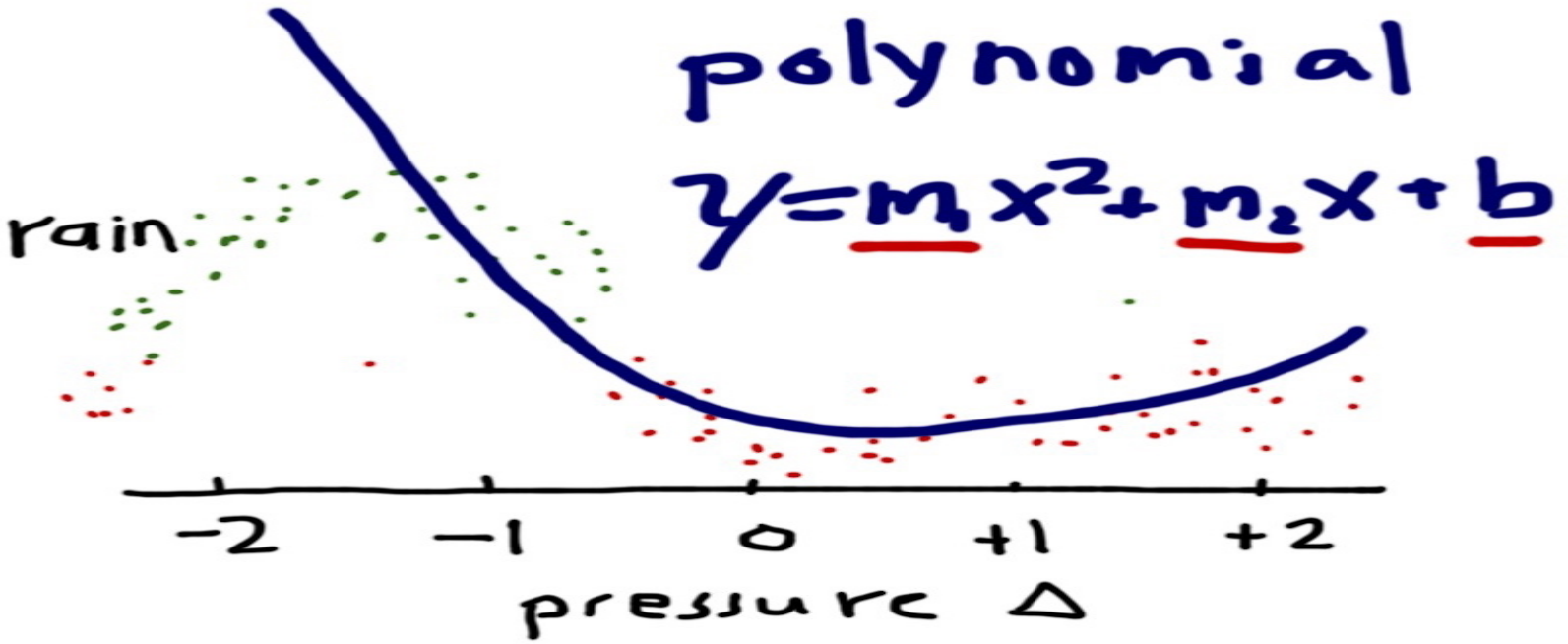
Example: Predict How Much Rain



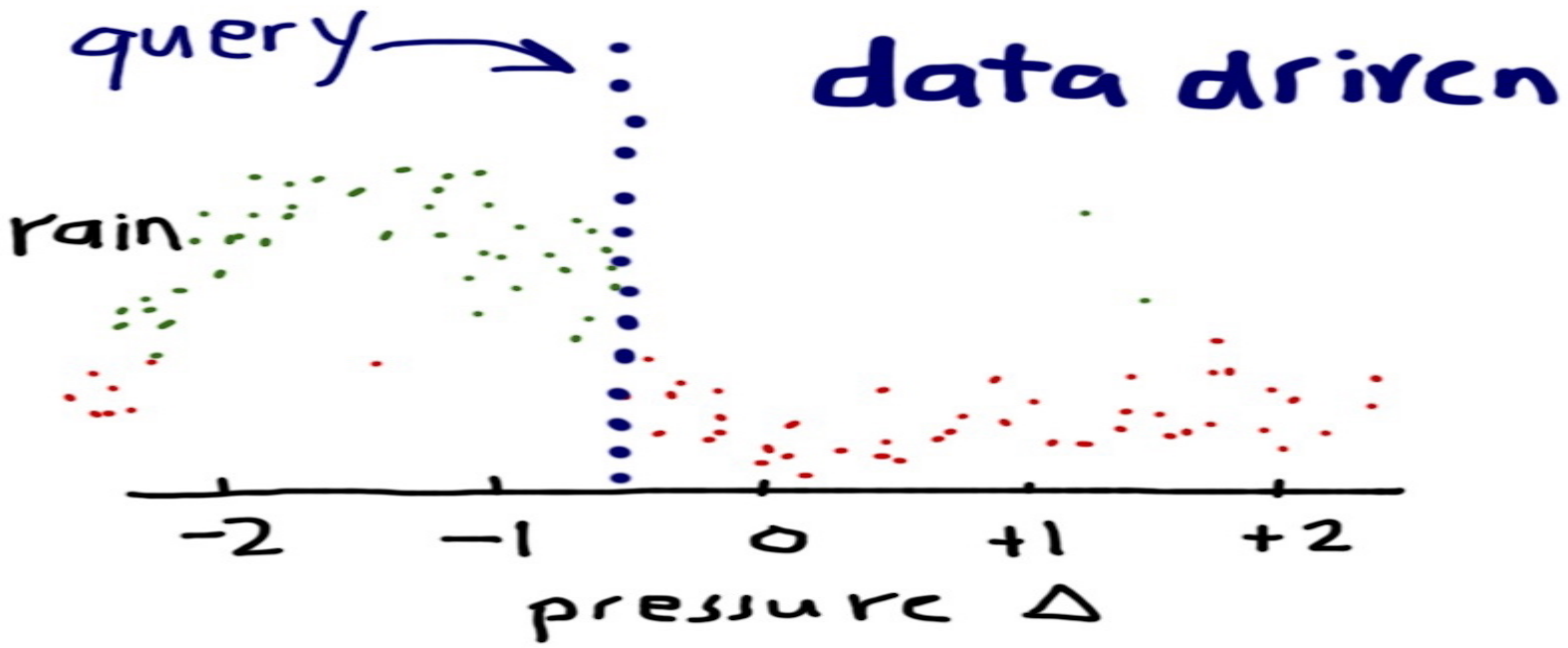
Solution 1: Parametric Model



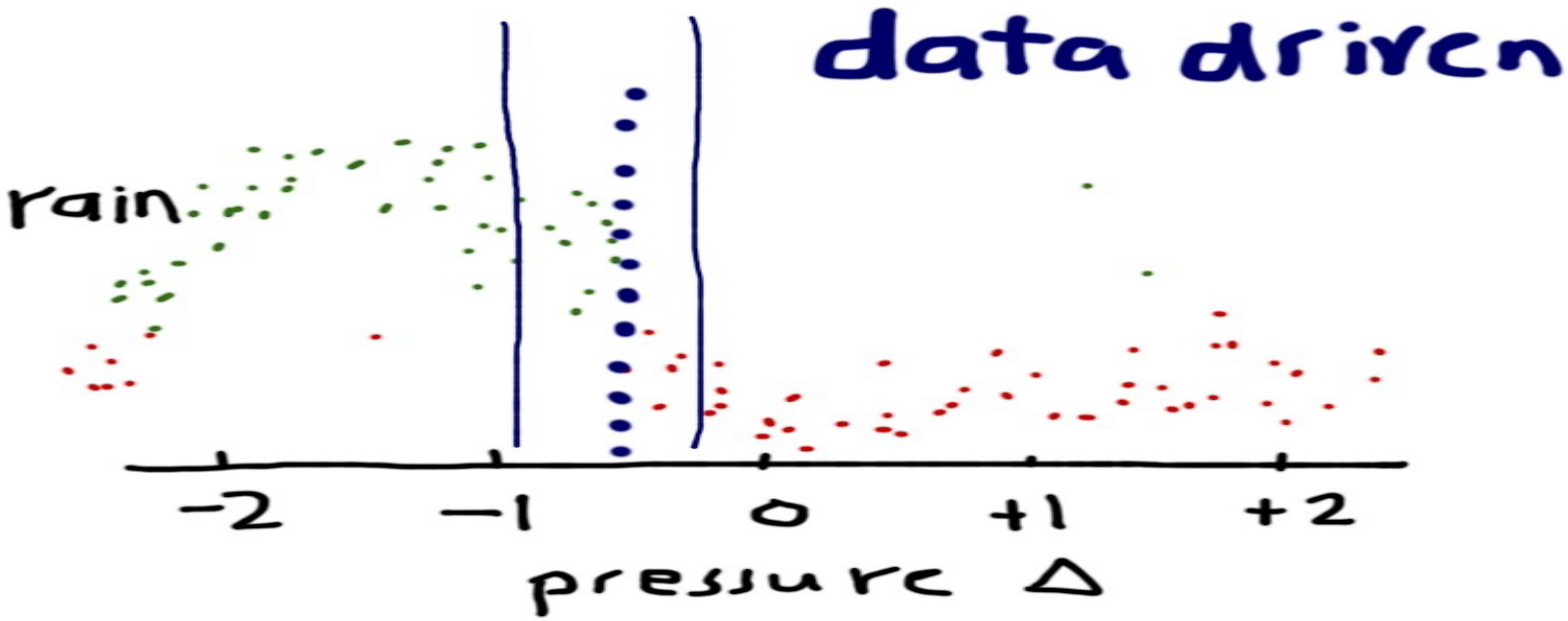
Solution 1: Parametric Model



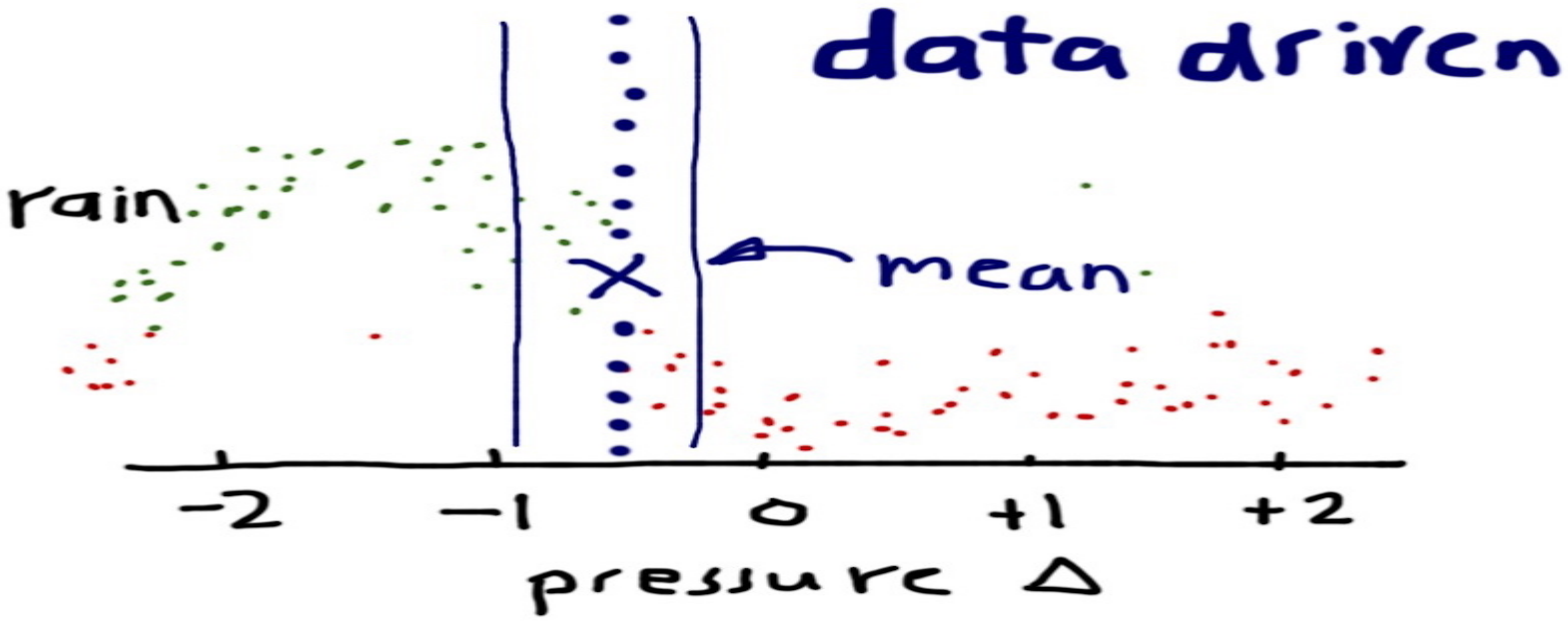
Solution 2: Data Driven



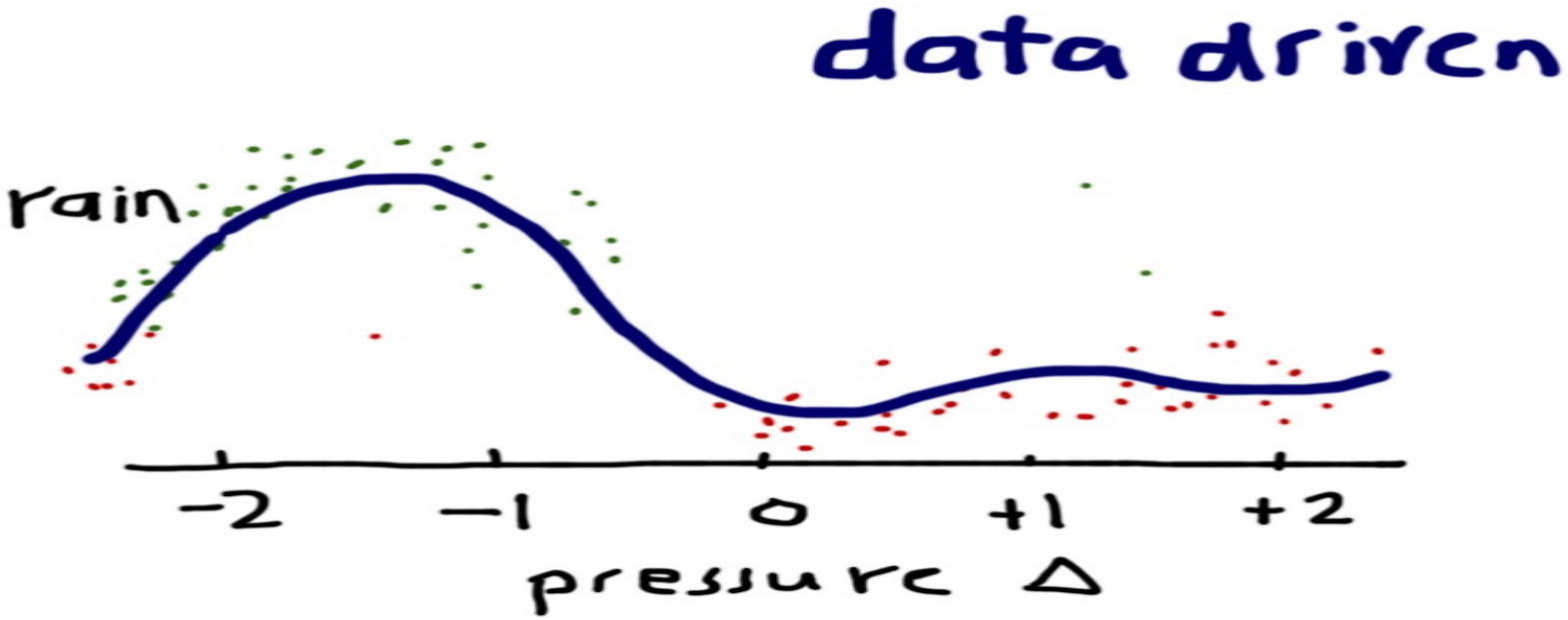
Solution 2: Data Driven



Solution 2: Data Driven

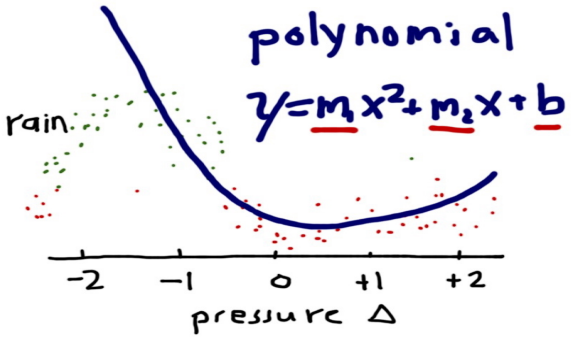


Solution 2: Data Driven

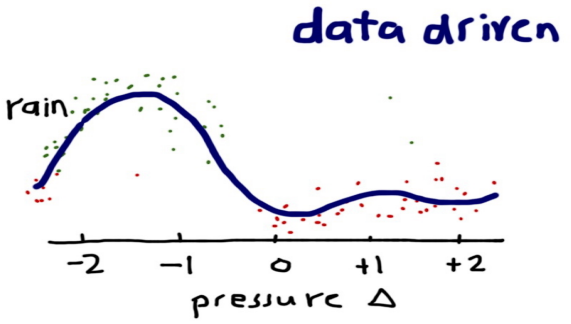


Data-Driven vs Parametric

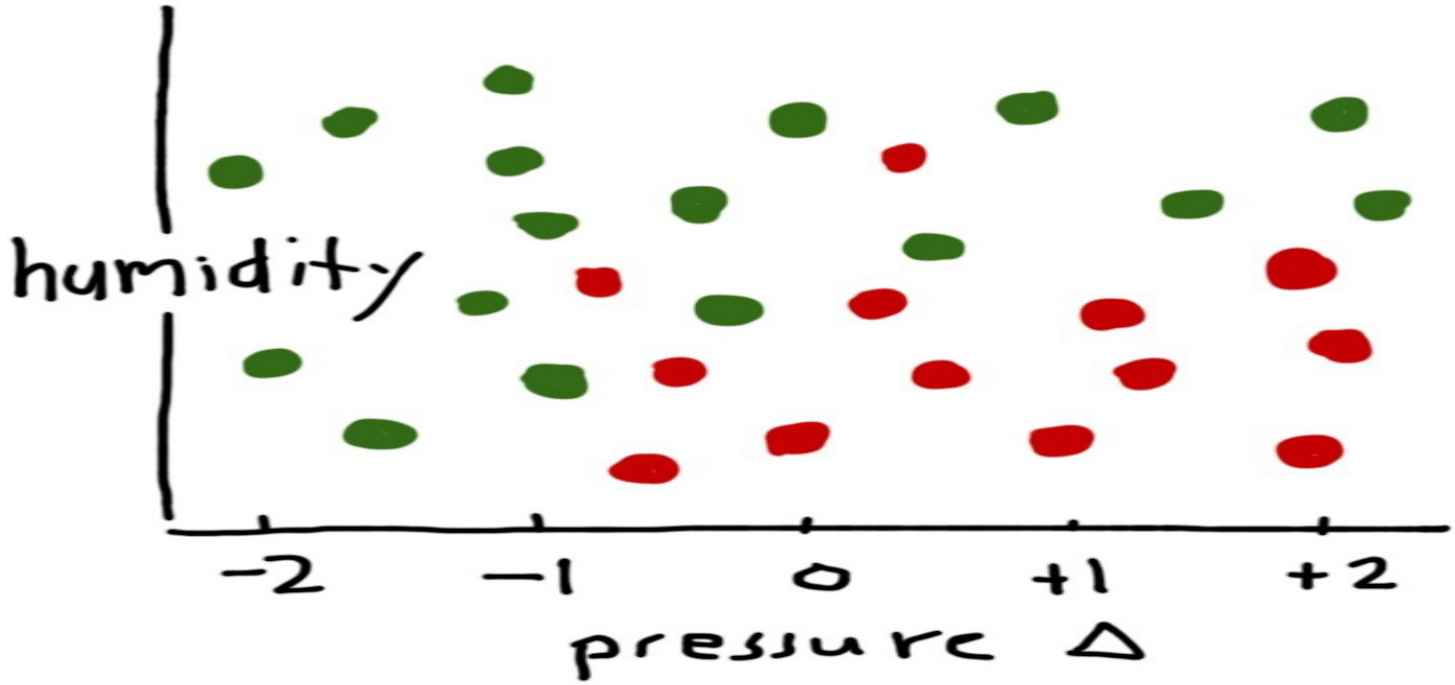
- Parametric Model Pros
  - Don't usually over fit
  - Very fast at run time
- Parametric Model Cons
  - Don't model complex non-linear data well
  - Sometimes over simplify



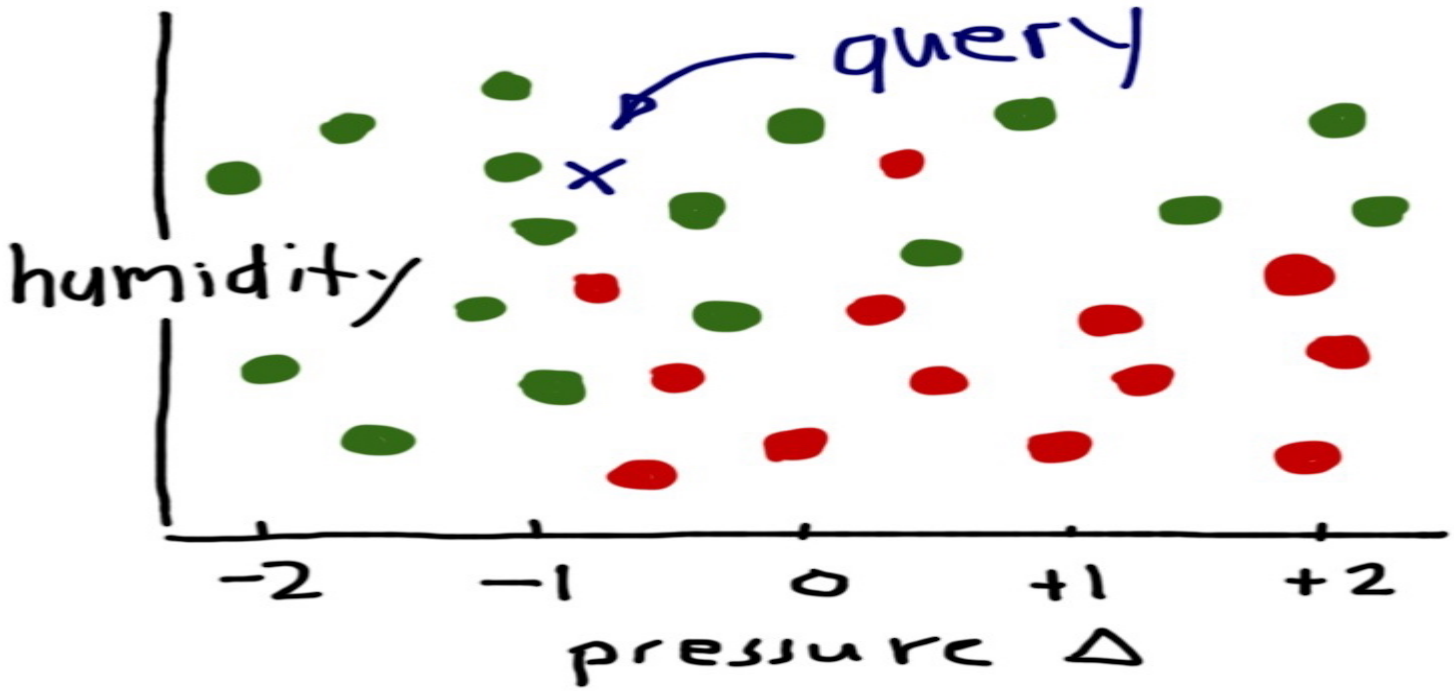
- Data-Driven Model Pros
  - Can model complex data well
  - Easily adapts to more data
- Data-Driven Model Cons
  - Sometimes slow at run time
  - Sometimes over fits



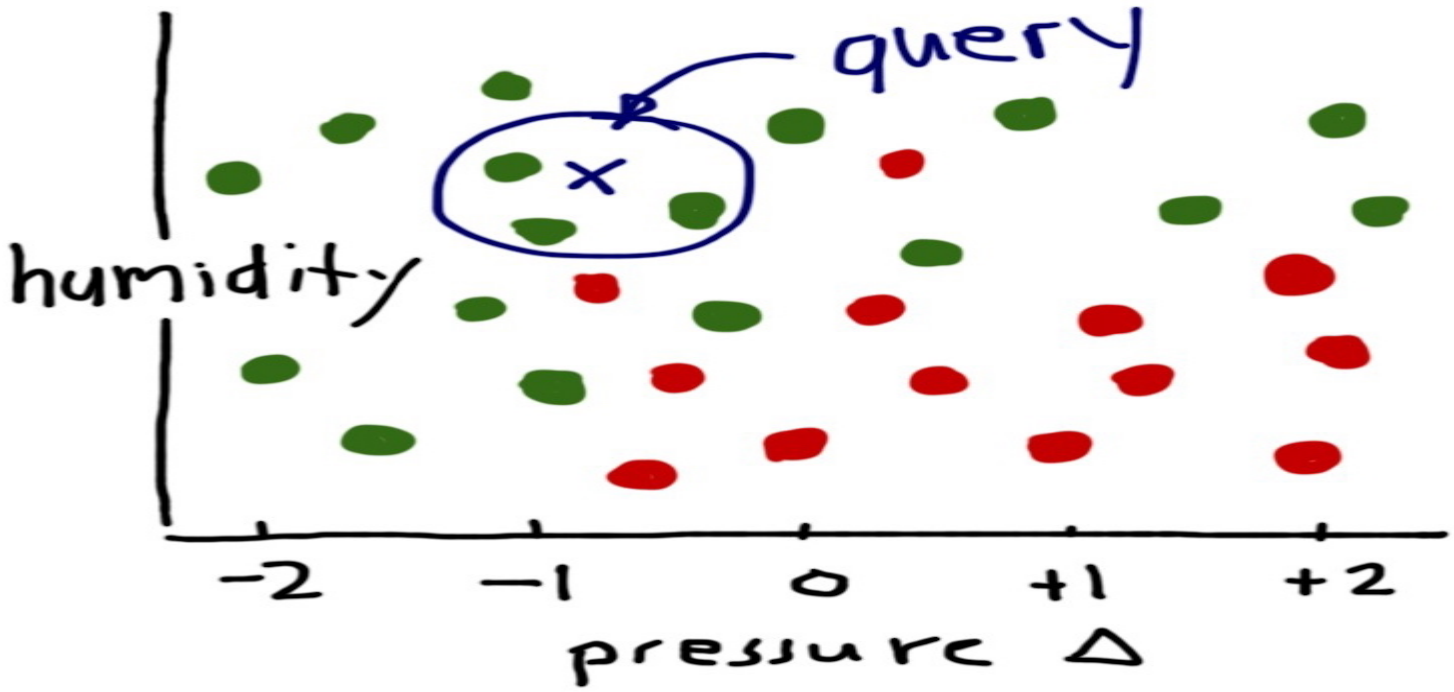
2D Data-Driven Model



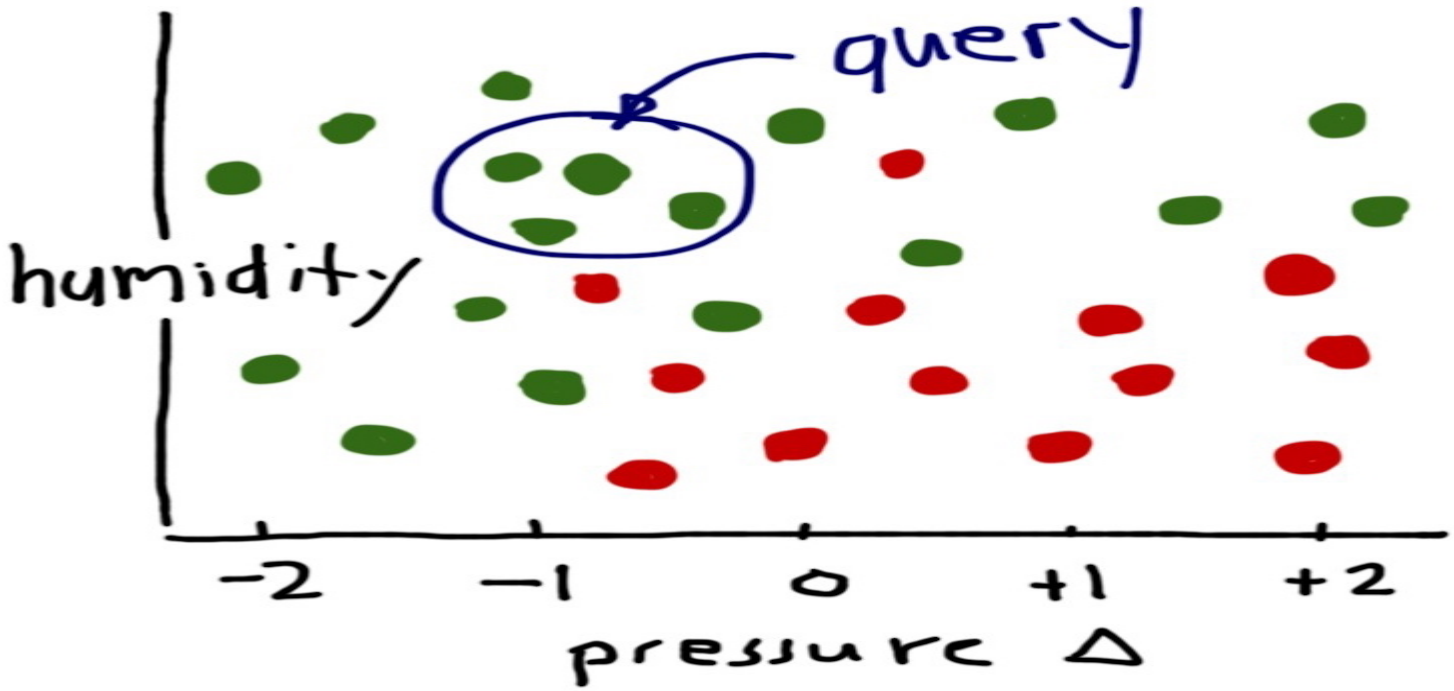
2D Data-Driven Model



2D Data-Driven Model



2D Data-Driven Model

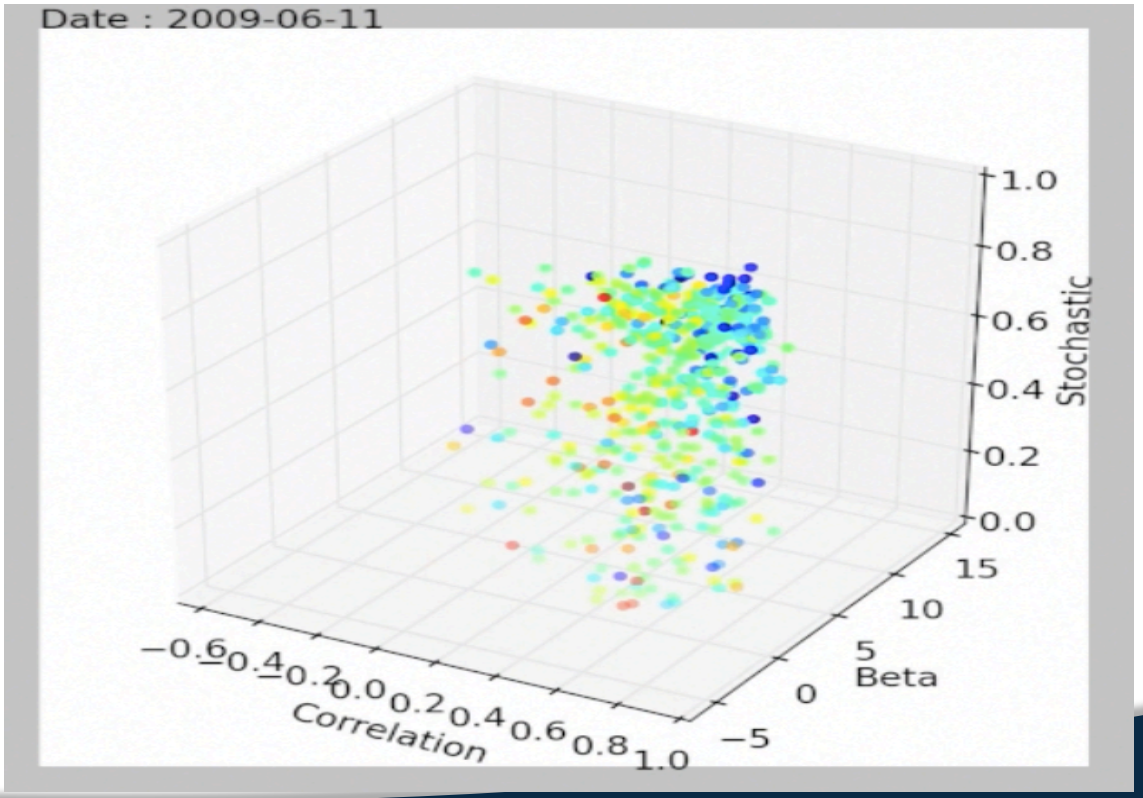


Application to Trading & Investing

- We can do the same thing with information about stocks

Weather	**Stocks**
Barometric pressure	P/E Ratio
Humidity	News
Rainfall	*Future Price*

Application to Trading & Investing



Other Considerations

- What are the best factors to use?
  - Your insight, or
  - Our number crunching?
- Which data to use for models?
- How to assess models?
- How to prevent over fitting?