Machine Learning

Balamurugan M

April 04, 2019

What is Machine Learning?

Machine learning is the practice of teaching computers how to learn patterns from data, often for making decisions or predictions

The Curious Child

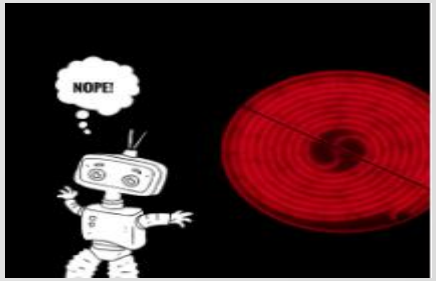
A young child is playing at home… And he sees a **candle**! He cautiously waddles over.

- Out of curiosity, he sticks his hand over the candle flame.
- "Ouch!," he yells, as he yanks his hand back.
- "Hmm… that **red and bright** thing really hurts!"

The Curious Child

Two days later, he's playing in the kitchen… And he sees a **stove-top**! Again, he cautiously waddles over.

- He's curious again, and he's thinking about sticking his hand over it.
- Suddenly, he notices that it's **red and bright**!
- "Ahh…" he thinks to himself, "not today!"
- He remembers that **red and bright** means pain, and he ignores the stove top.

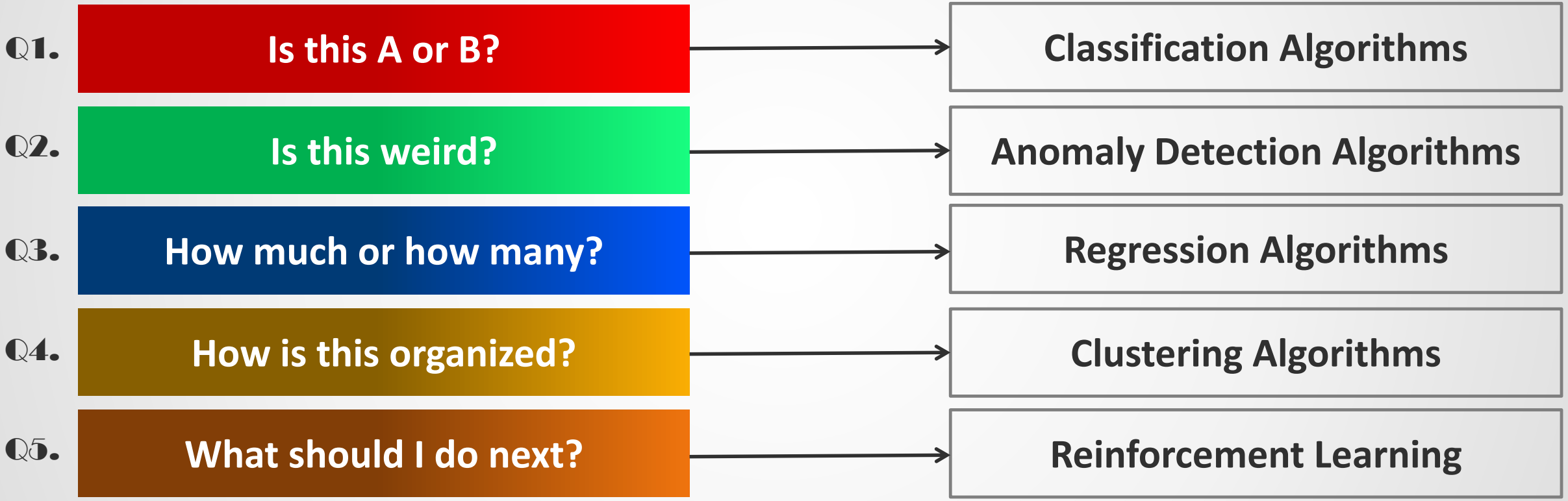


To be clear, it's only machine learning because the child learned **patterns** from the candle.

- He learned that the pattern of **"red and bright means pain"**
- On the other hand, if he ignored the stove-top simply *because his parents warned him*, that'd be "explicit programming" instead of machine learning

How a Problem is Solved Using ML?

- These are the 5 questions, which can be answered in Machine Learning:



These algorithms are fitted into three 3 ML categories. They are:

Machine Learning Types

1. **Supervised Learning**

Supervised learning is a type of machine learning algorithm that uses a known dataset (called the training dataset) to make predictions. The training dataset includes input data and response values. From it, the supervised learning algorithm seeks to build a model, which can make predictions for a new dataset

2. **Unsupervised Learning**

Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting data without labeled responses

3. **Reinforcement Learning**

Reinforcement learning is an important type of Machine Learning where an agent learn how to behave in an environment by performing actions and seeing the results

Machine Learning Types

Machine Learning

Supervised

Regression
- Linear
- Polynomial

Decision Tree

Random Forest

Classification
- kNN
- Logistic Regression
- Naïve-Bayes
- SVM
- CART

Unsupervised

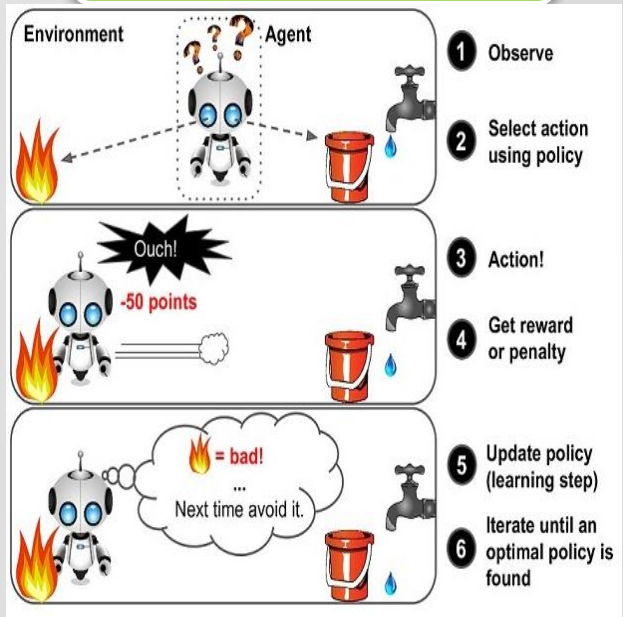
Clustering
- Agglomerative Hierarchical Clustering
- k-Means

Continuous

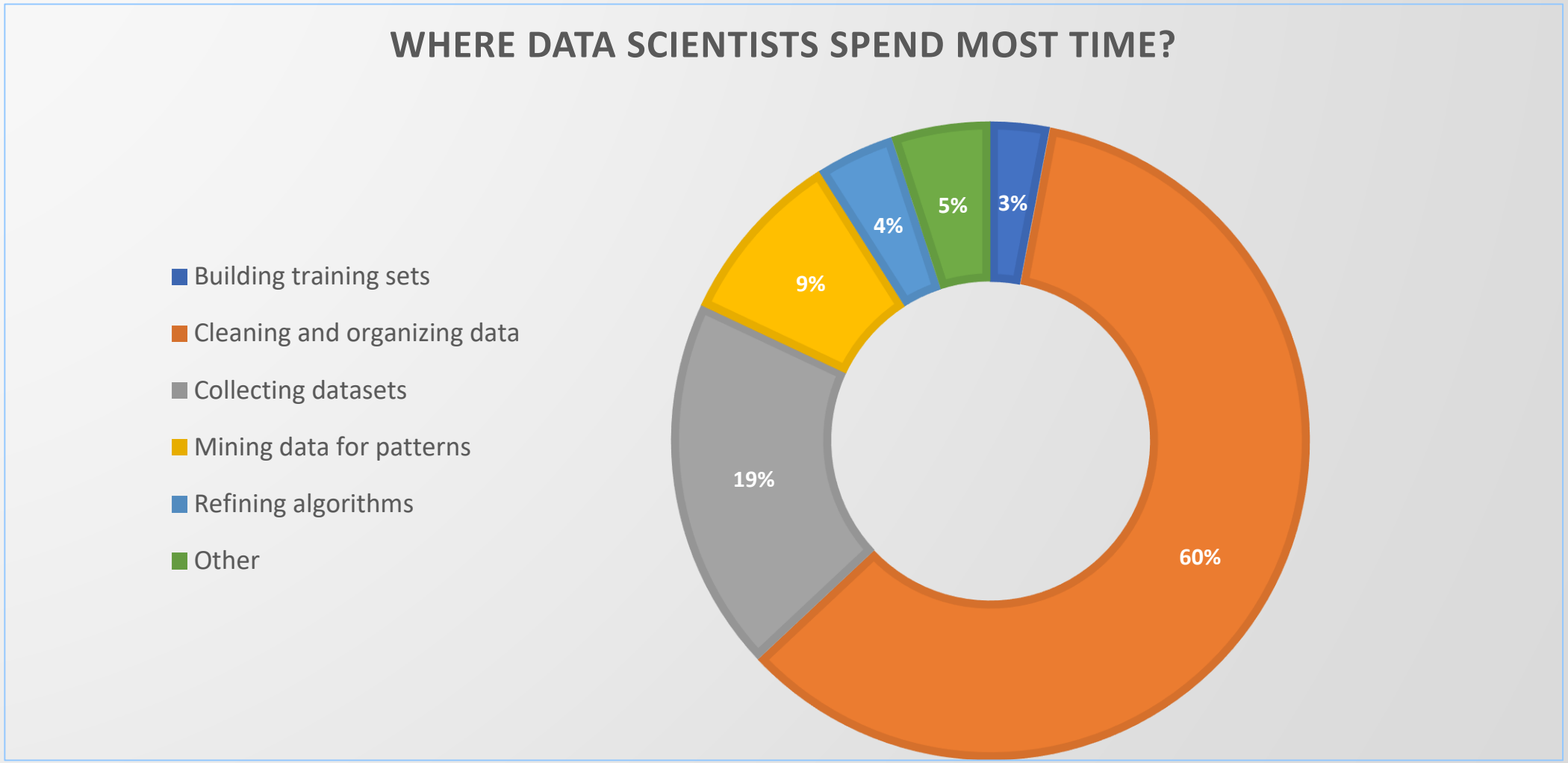
Association Analysis
- Apriori

Categorical

Reinforcement



What Goes into a Successful Model?



Where ML Algorithms are Used?

Regression Algorithms

- Regression Algorithms are used to calculate/predict numerical values

Use Cases

- Predicting Credit Limit for a new Credit Card?
- What will be the temperature tomorrow?
- How much discount can you give on a particular item?

Classification Algorithms

- Classification Algorithms are used to classify a record
- Binomial Classification – only 2 choices
- Multinomial Classification – More than 2 choices

Use Cases

- Email Classification – Spam or Ham?
- Will he/she buy the product?
- Churn Prediction

Where ML Algorithms are Used?

Clustering Algorithms

- Clustering separates the data into groups or clusters, to ease out the interpretation of the data

Use Cases

- Market Segmentation?
- RFM Analysis for Customer Segmentation

Anomaly Detection

- It analyzes a certain pattern and alerts you whenever this is change in the pattern

Use Cases

- In real life, your bank uses Anomaly Detection algorithms, and flag any transaction, which is not usual as per your transaction

Association Rules

- Association Rule Mining is one of the ways to find patterns in data. It finds:
  - features (dimensions) which occur together
  - features (dimensions) which are “correlated”

Use Cases

- Market Basket Analysis

How Do You Start?

Prerequisites?

Basic Statistics & Programming

Preferred Languages for Data Science

- Python
- R
- MATLAB
- Julia
- Java
- TensorFlow

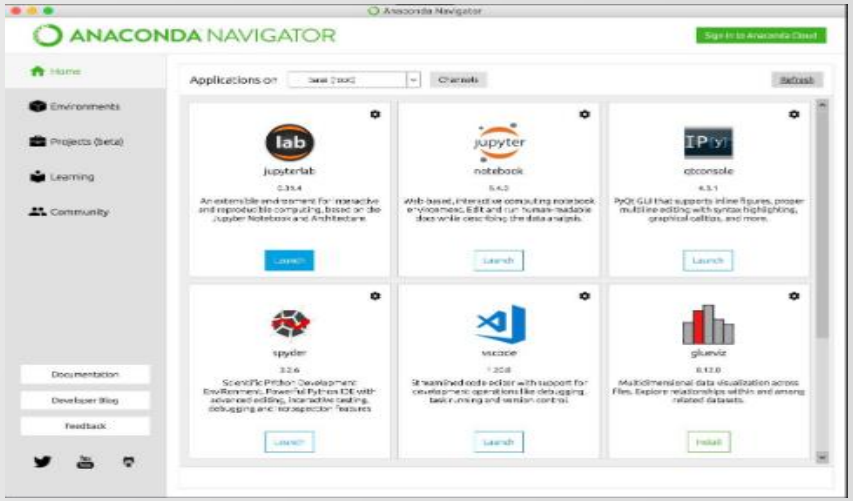
Where do I get trained? (Udemy - Hexaware)

https://hexaware.udemy.com/organization/home/courses/ufb-data-science/

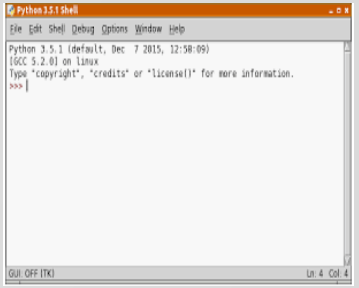
7 Best Data Science Courses & Certifications for 2019:

1. Data Science Specialization — JHU (Coursera)
2. Introduction to Data Science — Metis
3. Applied Data Science with Python Specialization — UMich (Coursera)
4. DataCamp
5. Statistics and Data Science MicroMasters — MIT (edX)
6. CS109 Data Science — Harvard
7. Python for Data Science and Machine Learning Bootcamp — Udemy

Where to Code?



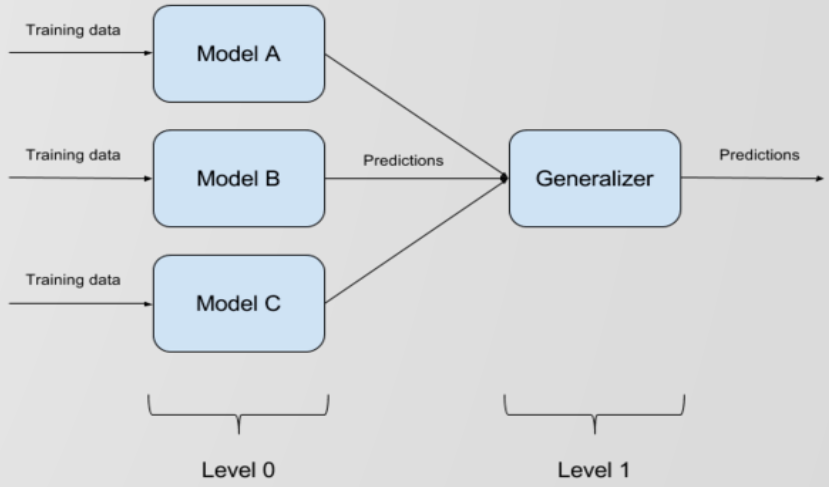
Microsoft Azure Notebooks



What Next?

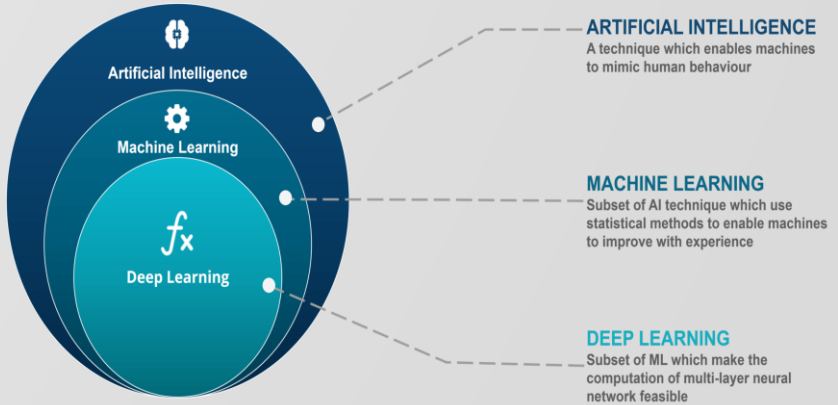
1. **Ensemble Machine Learning**

Ensemble Machine Learning methods combine several machine learning algorithms and techniques into one predictive model in order to decrease the variance (bagging), bias (boosting) or improve the predictions.

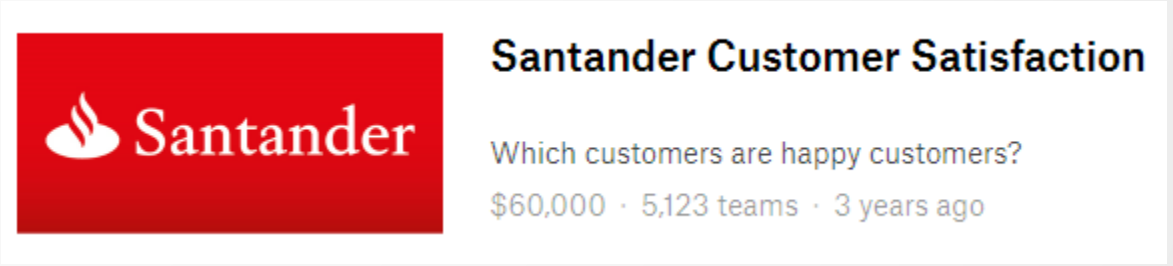


2. **Deep Learning**

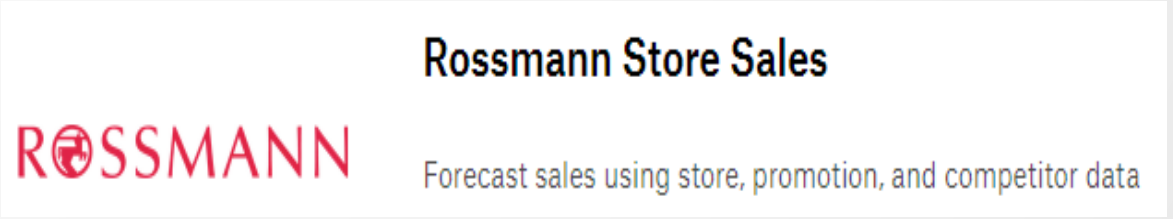
Deep learning is a machine learning technique that teaches computers to learn by example



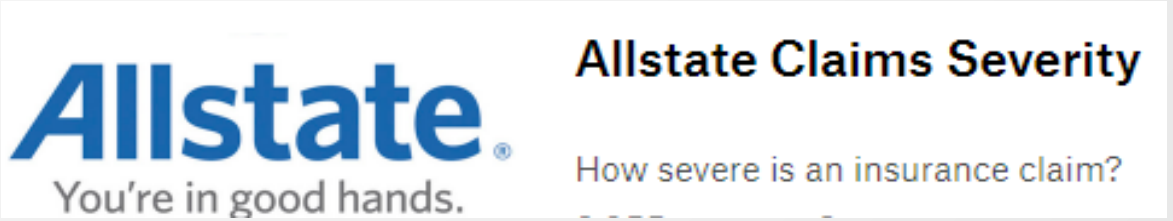
A Few Industry Challenges...



Santander Bank is asking ML experts to help them identify dissatisfied customers early in their relationship. Doing so would allow Santander to take proactive steps to improve a customer's happiness before it's too late.



Rossmann store managers are tasked with predicting their daily sales for up to six weeks in advance. Store sales are influenced by many factors, including promotions, competition, school and state holidays, seasonality, and locality.



- **All State** wants to predict the severity of an insurance claim
- Predict the likelihood and cost of insurance claims