Section 1 AI Fundamentals

Monday, April 5, 2021 4:14 PM

Lesson2: What is AI?

- **Artificial Intelligence** - The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.
- **Artificial General Intelligence** - That robot or computer in the movies that can teach itself a new task. DOES NOT CURRENNTLY EXIST TODAY.
- **Narrow Artificial Intelligence** - Computer systems which use human intelligence but have stong limitations in what they can do. Ex. Alexa, Siri
- **Machine Learning** - The study of computer algorithms that improve automatically through experience.
    - **Unsupervised Learning** - Ability to find patterns in data without human help.
    - **Supervised Learning** - Humans label the data and give general guidance.
- **Natural Language Processing (NLP)** - Allows a machine to read and understand human language.
    - Machine translation, question answering, sentiment analysis, etc.
- **Perception** - The ability to use input from sensors - images, audio, lidar, sonar, radar, touch.

Lesson 3: What is an ML Model?

- **Machine Learning** allows computers to use data to forecast the future, without specifically being programmed.
    - A machine learning algorithm would need to analyze traffic patterns and data to decide what the best outcome or suggestion would be for the given circumstance.
- IN ML, a **Model** is a program that can be used to recognize a pattern in data.
    - Ex. You would use a model to examine traffic lights and patterns to see the times lights needs to be a certain color.
    - A model can be used to predict future behaviors.
    - A model can be used to categorize something as on things or another.
    - A model can be used to recognize people, objects and landmarks using unseen images.
    - A model can be used to understand the context of natural human text or speech.
- You train a model using "training data".
- You evaluate a model using "test data" to measure how accurate it is.
- Once a model has been deployed, it can recognize patterns in data it has never seen before.

Section 2 Artificial Intelligence workloads and Considerations

Monday, April 5, 2021 4:34 PM

Lesson 6: Common AI Workloads

1. **Prediction and Demand Forecasting**

- Using machine Learning to Predict **(Supervised Learning)**
  - Give the machine all the relevant data you know.
  - Tell it which field you want to predict.
  - It develops a model which it uses to make a prediction.
  - EX. Google Searches are predictive models to predict search terms. Same for Netflix and how the select which shows to provide to you as suggestions.

2. **Anomaly Detection**

- Ability to look at previous data and detect anomalies and outliers.

3. **Computer Vision**

- Ability to use vision tools such as a camera, to identify and analyze an image.
  - Ex. Be able to tell in a picture someone is swimming, or how old someone may be.

4. **Natural Language Processing (NLP)**

- Ability for the computer to infer if someone is asking a question, and deals with speech or text. Ex. Ask Siri

5. **Conversational AI - Chat Bots**

- AI bots that are used on websites to help answer questions, and it will recognize what the user is asking to give an appropriate answer.

Knowledge Mining - Cognitive Speech

- Ingest content from Blob storage, tables, databases, etc.
- PDF's, images, word documents, etc.
- Uses AI tools such as image classification, face recognition, etc, extraction to create a searchable index in the clutter.

Lesson 7: Guiding Principles for Responsible AI

- The prevalence of AI causes some ethical and moral challenges.
- **Unintended Consequences** - Computer could make decisions that are wrong, illegal, can't be explained, or harmful to society at large. We want to prevent this so these must be created with DD in mind. When algorithms are making decisions on what to show people, things can go wrong.
- Six Principles should guide AI Development
  - Fairness
  - Reliability and Safety
  - Privacy and Security
  - Inclusiveness
  - Transparency
  - Accountability

Lesson 8: Fairness, Reliability & Safety and Privacy & Security

- **Principle of Fairness** - AI Systems should treat everyone fairly and avoid affecting similarly situated groups of people in different ways.
  - Ex. AI system that prioritizes patients in the ER. Some other factor that isn't related to their health may give them priority. This is unfair. Sees how much someone's net worth is and gives people with more money priority.
- **Principle of Reliability and Safety** - To build trust, its critical that AI systems operate reliably, safely, and consistently under normal circumstances and in unexpected conditions.
  - It's so easy to program for the "sunny day" where nothing ever goes bad and things will always work. Testing and PenTesting is important so you can find weaknesses and code for

inaccuracies.
    - Ex. How does a self driving car work not only on a sunny day, but in the snow, with no road markings, etc.
- **Principle of Privacy and Security** - Many countries and regions in the world are developing new standards and laws to try to protect its citizens. Laws are always slower than technology.
    - Ex. Target knowing a women was pregnant before her dad based on advertisements and laws around these kind of things.

Lesson 9: Inclusiveness, Transparency, and Accountability

- **Principle of Inclusiveness** - Everyone should benefit from intelligent technology, meaning it must incorporate and address a broad range of human needs and experiences.
    - How does stuff work when challenges are present?
    - Ex. How does voice assistant work for someone who has a speech impediment
    - Ex. How does a "video-only" product instructions help someone who is blind.
- **Principle of Transparency** - When AI Systems are used to help inform decisions that have tremendous impacts on people's lives, it is critical that people understand how those decisions were made.
    - Ex. Rejected from an AI system for a job, should be able to explain why they were rejected. If you do not know, there is no way to tell if it is unfair or if it made an incorrect decision.
- **Principle of Accountability** - The people who design and deploy AI systems must be accountable for how their systems operate.
    - There should be regular review of how AI is operating and regular improvement of the model. This way AI systems stay in-line up to date, and make these systems accountable and reliable.

Section 3 Identify Common Machine Learning Types and Core Machine Learning Types

Monday, April 5, 2021 5:18 PM

Lesson 11: Common Machine Learning Types

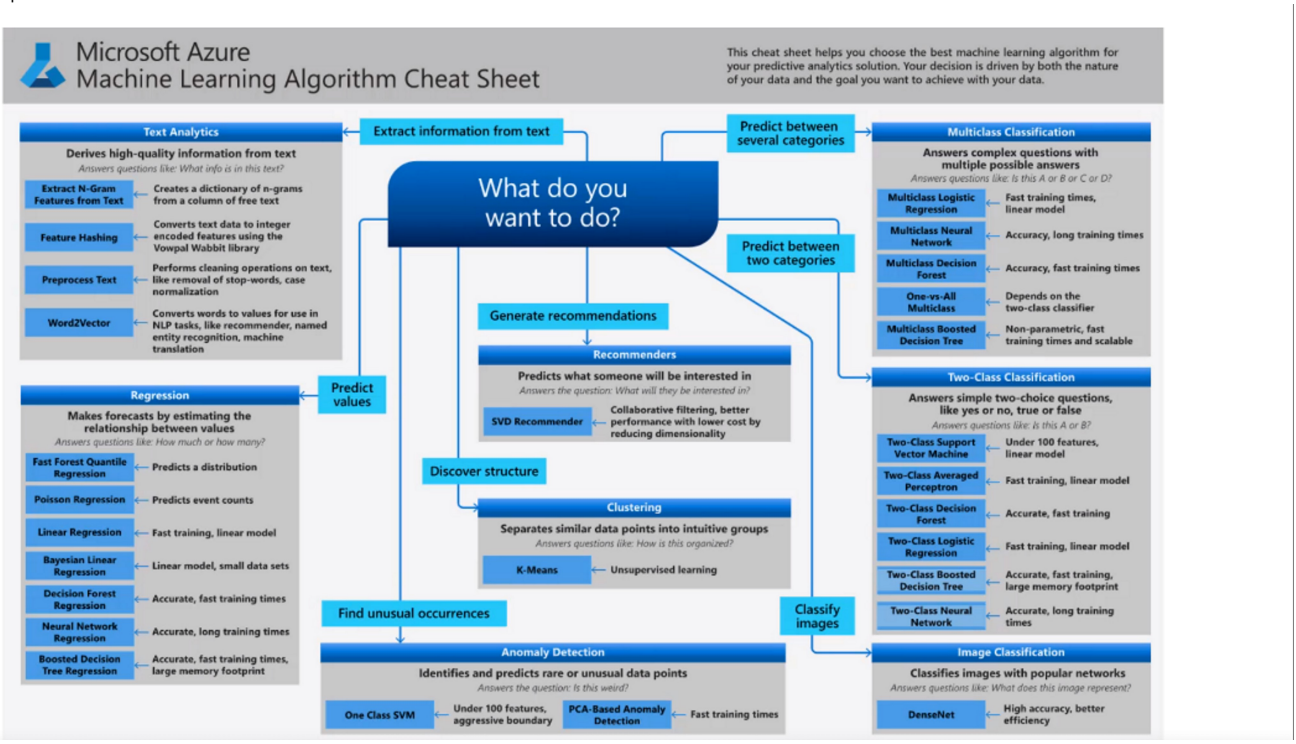
- **Regression** - A type of supervised learning. The ability to predict the outcome variable given 1 or more predictor variables.
  - **Result is numeric - price, amount, size, etc.**
  - Finds the relationship between the variables, between the X and Y axis.
- **Classification** - A type of supervised learning.
  - **Cluster Analysis** - Assign a score to the odds of it belonging to a cluster.
  - Asking computer what kind of fruit this is?
  - **Binary Classification** only has two answers, 0 and 1. Ex. It is or is not an apple, yes or no.
  - **Multi-class Classifications** allow for other options. Ex. Knows what an apple, orange, banana, etc. look like and will assign percentages based upon that info.
- **Clustering** - A type of unsupervised learning. Find groups of related things among data on its own without labeling data manually. Will find information generating insight you did not even know.
  - Ex. What do my customers have in common that allow me to go find more customers.

Lesson 12: Dataset Features and Labels

- **Feature** is an input Variable. When you are marking the data that an ML algorithm is going to use to make its decisions, that is a feature. This is the input into the algorithm.
- **Label** is the thing we're predicting.
- Given a pile of data, you (data scientist) need to determine which bits are relevant to make decisions on.
  - Examples of Data: First and last name (feature), smoker or non-smoker (feature), DOB (feature), gender (feature), income (feature), # of children (feature). Do these data types help to figure out Life Expectancy (Label).

Lesson 13: Training and Validation Datasets

- **Training the Model**: The more input (historical) data you have, the more accurate the results. Don't use all your data to train the model.
  - Divide your available data into Training and Validation/test datasets.
  - Ex. 1,000,000 rows, split 500,000 to training and 500,000 to validation.
  - Split the training and validation datasets randomly.
- **Evaluate the Results - Regression**: Use the validation datasets to test the model, and measure how close or far the actual results are from the predicted results.
  - **Mean Square Error** - Large differences are much worse than small differences.
- **Evaluate the Results - Classification**: The result is to give a prediction score that the subject is part of the group. Ex. "70% confident this is an apple, 30% confident this is a pear."
  - **False Positives vs False Negatives**: Compare true positives with false positives and true negatives with false negative when evaluating the model. How important is it to you that it never have a false positive.
    - Accuracy vs. precision

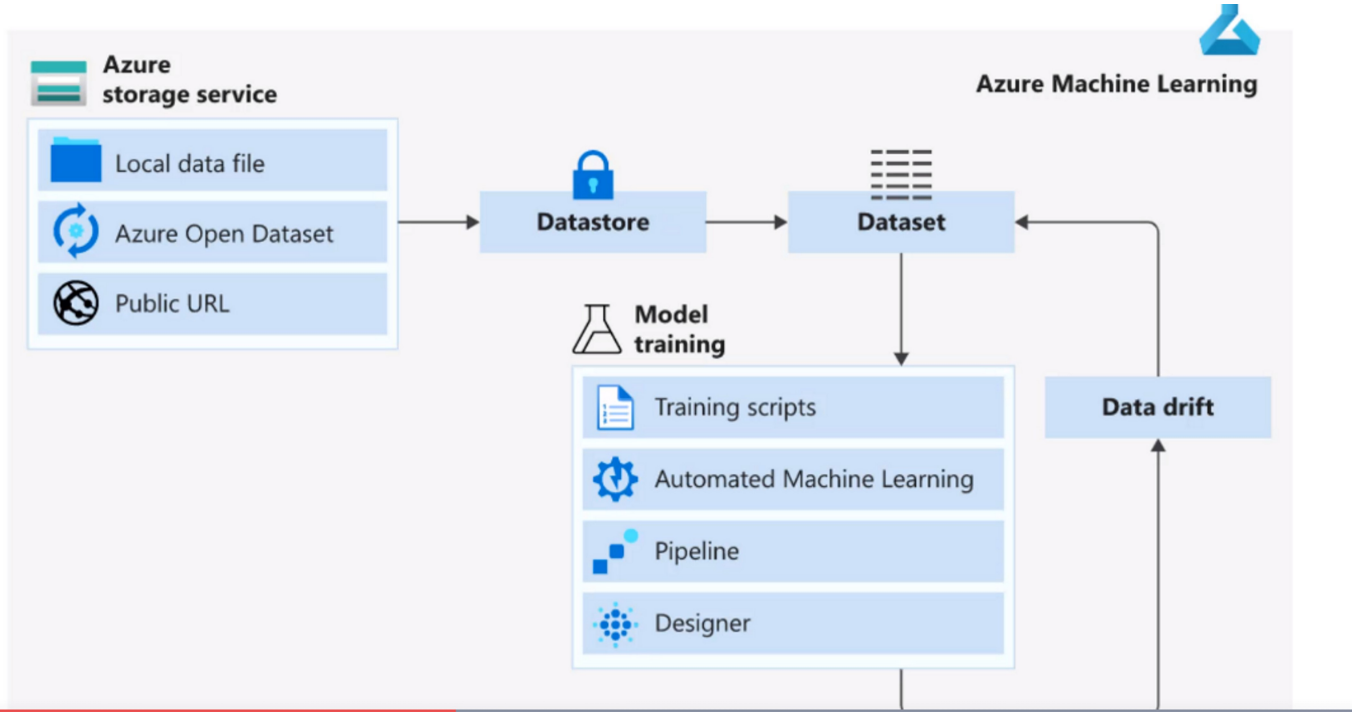


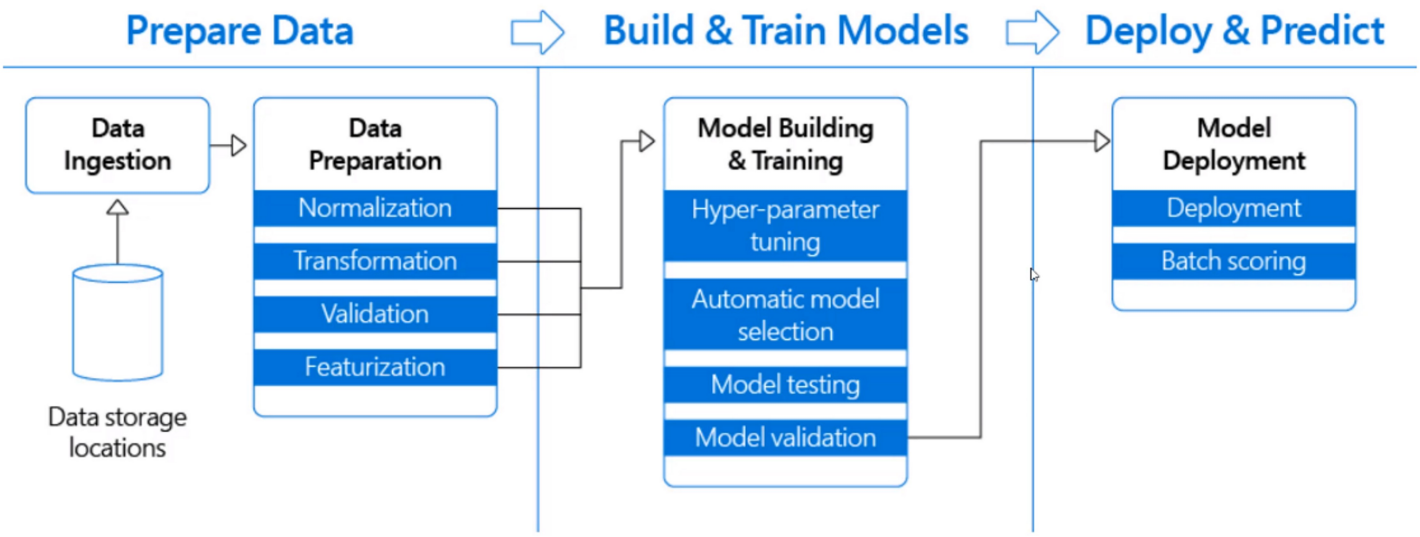
Section 4: Core Tasks in a Machine Learning Solution

Monday, April 5, 2021 5:48 PM

Lesson 14: Data Integration and Preparation

- **Azure Data Factory** is a data extraction tool in Azure, typically referred to as ETL (Extract, Transform, Load).
  - Raw data, pull the data, transform the data, you get prepared data, then train the model.
- **Data Store** is a connection string to your data. Then you use your **Data Set** to then release into your model training.
- **Steps to Analyzing Data**
  - Load your data into Azure Storage from some other source.
  - Clean up the data to remove low quality / bad data
  - Transform the data into useful / standard formats,. Intelligent defaults.
  - Create a datastore in Azure ML to link to the data.
  - Create a dataset in Azure ML to work with.
    - File or table data.
  - Register the dataset in your workplace.
- Data Ingestion Tools
  - **Azure Data Factory** - The ETL solution.
  - **Azure ML Python SDK** - Scripts and custom code.
- Data Processing
  - Parquet and CSV file formats
  - Pandas dataframes for data less than 1GB
  - Ray, Dask, or Spark for data more than 1GB





Lesson 15: Feature Engineering and Selection

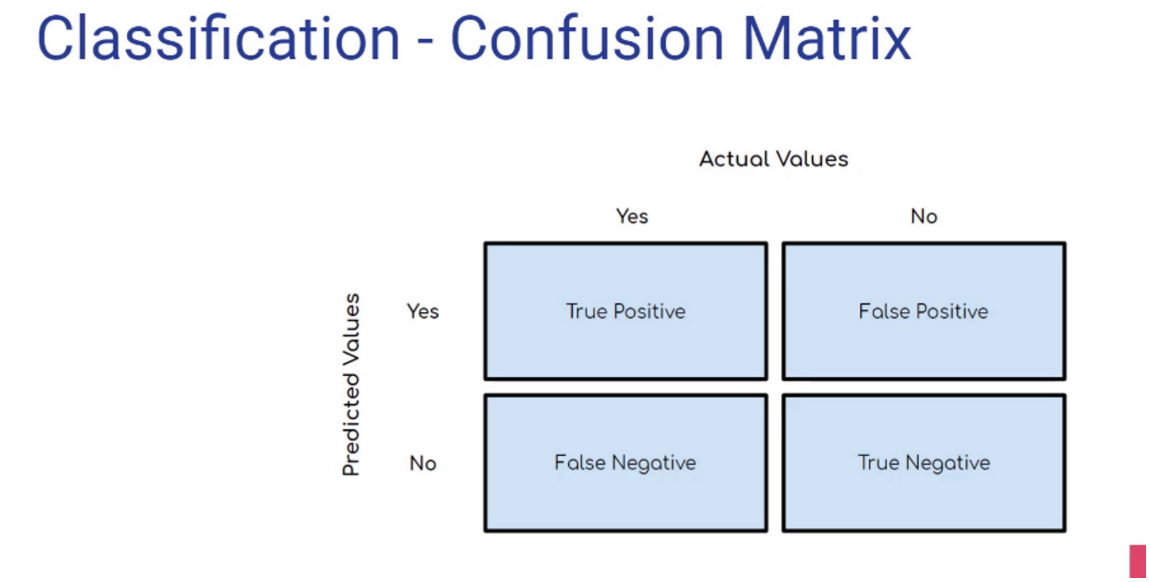
- **Feature Selection** - The process of selecting a subset of relevant features for use in model construction.
  - Azure ML supports both automated and manual configuration of ML.
  - **Feature Selection** is one of the core concepts and has a huge impact on the performance of the model in ML. You need to identify and remove irrelevant or partially relevant features.
- **Over-fitting the training data** - This is what happens when you select to many features, meaning that your model can only predict data that it has already seen, resulting in huge errors when seeing data not exposed to the system before. This is an over-trained model.
- **Simplify the Model** - Feed in less features, resulting in a more simple model, bringing probably a better result.
- **Minimum Redundancy, Maximum Relevance:** a Redundant field could be a city name and phone area code, as there is a high correlation between the two so using both would overfit the model and make it do to much work.
- **Curse of Dimensionality** - Similar to overfitting. Gives no predictive power because of unique fields. LOOK THIS UP FOR MORE DETAIL!
- **Feature Engineering** - The process of using domain knowledge to extract features from raw data via data mining techniques. Construct new columns based off existing columns or another source, using the engineering aspect to create new data based off existing data.
  - Ex. Turn their date of birth column into the age column.
- Brainstorming and iterating on the best features to achieve the desired result.

Lesson 16: Training and Evaluation

- Training a model is providing an ML algorithm with training data to learn from.
- **AutoML** - Try all the training algorithms to see what one is the best.
- **ML Designer** - You choose the algorithm to use.
- **Model Training** is…
  - The process by which a ML Algorithm "learns"
  - Depending on the type of ML (Regression, Classification, clustering)
  - Feed it your training dataset.
  - Could take hours to process
  - May or may not be correct until you evaluate it.
- **Model Evaluation** is…
  - The process by which you estimate the accuracy of the model
  - You typically use a test dataset which is different than the training dataset.
  - Compare the prediction made by the model with actual result that you already know.
  - There are various ways to determine what is a good result and what is a bad result.

Lesson 17: Evaluation a Classification Model: The Confusion Matrix

- **Classification - Recall**: Recall = TP / (TP + FN): Out of all positive outcomes, how many did your predict correctly? Should be as high as possible.
- **Classification - Precision**: Precision = TP / (TP + FP): Out of all positive predictions, how many were actually positive. The higher the number the better.
- **Classification - Accuracy**: Accuracy = (TP + TN) / Total: How many correct predictions did you get out of the total predictions?
- **F-Score** combines recall and precision into a single score: F-Score = 2 *recall *precision / (recall + precision)



Lesson 18: Evaluating a Regression Model

Section 5: No Code Machine Learning

Tuesday, April 6, 2021 8:42 AM

Lesson 20: Auto ML

- **AutoML: Automated Machine Learning**
  - Machine learning with no code involved usually using visual tools.
- **Azure Machine Learning Designer**
  - Drag the dataset on to the designer, visualize it, and exclude columns that are not going to be useful for training your model. Can clean rows with missing data, normalize the data to make it more useful through transformations, and drag training models onto the canvas.

Section 6: Computer Vision Workloads on Azure

Tuesday, April 6, 2021 10:47 AM

Lesson 25: Common Types of Computer Vision Workloads

- **Image Classification** - You can feed an image to that service, and it will come back with a single term for what that object may be. Ex. car
- **Object Detection** - This can identify multiple objects present in an image. Ex. Maple Leaf, can find Bicycle and the wheels.
- **Semantic Segmentation** - Identifying the boundaries of objects. Input image can be introduced such as houses, and it will make a drawing showing the edges of the houses.
- **Optical Character Recognition (OCR)** - This is able to read text on images and put it into digital format.
- **Facial Detection and Recognition** - Can detect celebrities and faces to determine who a person may be.

Lesson 26: Computer Vision Services in Microsoft Azure

- **Cognitive Services** - Includes many services under one umbrella. All services require a single endpoint and a single key.
- **Computer Vision Services**:
  - Pre-trained ML model
  - Can recognize over 10,000 objects
  - Can generate automatic captions for images and tags
  - Content moderation for adults
  - Detect Faces
  - Text Recognition
- **Custom Vision Service**
  - A model that you can build and train
  - Classification or object detection
  - Upload an existing data set of images and classes
- Custom vision service separates are training and prediction into two resources.
- **Face Service** - Can recognize a human face in an image
  - Returns the rectangle coordinates of those 1 or more faces.
  - Need to be trained on your own data.
  - Can detect celebrities.
- **Face Detection** can detect age and gender using face services.
- **Face Verification** - To compare one face against a known face to identify if they belong to the same person.
- **Similar faces** - Identify faces that are similar to that person.
- **Facial Grouping** - Send in a set of unknown faces and divides them into groups based on similarity. Also feat. Differentiation based on expression.
- **Identify API** - Put a name to a picture so when it sees a new picture of someone it is trained to recognize, it can put the name with it.
- **Form Recognizer** - This can identify various fields on forms and extract the data. Ex. Invoices and receipts

Section 7 Natural Language Workloads on Azure

Tuesday, April 6, 2021 11:01 AM

Lesson 29: Common Types of NLP Workloads

- NLP - A computer can understand written and spoken language.
- Key Phrase Extraction - Identify the main points of a document and it understands the concept of that. Extracts main points of conversations to get the context out of a conversation.
- Entity Recognition - Identifying "entities" of a document; items categorized by type and subtype.
    - Ex. Snow, driveway, finger - objects
    - Earlier today - date
    - Driveway - location
- Sentiment Analysis - Sentiment score from 0 to 1 - 1 being positive sentiment.
- Language Modeling - Build your own dictionary for terms in your industry. Train the language understanding engine for your own use.
- Speech Recognition and Synthesis -
    - Speech Recognition - The ability to detect and interpret human speech.
    - Speech Synthesis - The ability to generate spoken output.
    - Speech to text and text to speech
- Translation - Supports over 60 languages, understands semantic concepts. Availability to use for different languages from lets say, English to Spanish. Closed Captioning is an example.

Lesson 30: NLP Services in Microsoft Azure

- **Text Analytics Service** - All of the azure services that include key phrase extraction, entity detection, and sentiment analysis.
    - It can detect language of text.
    - Can detect when text has multiple languages and identify the predominant one.
    - Returns Nan when it can't determine.
- **Language Understanding Service (LUIS)**
    - Understanding natural language using three core concepts.
        - **Utterances** -Something that you say such as "when do you close?"
        - **Entities** - The entities will then be analyzed. "Today" = date/time
        - **Intents** - The purpose or goal expressed by the user. Things your application is able to do
            - Goal: TodayHoursOfOperation

Section 8 Types of Conversational AI Workloads

Tuesday, April 6, 2021 11:15 AM

Lesson 31: Common Types of Conversational AI Workloads

- **Conversational AI** - Allows a human and a computer to talk.
  - Typically called Agents or Bots that can communicate over voice, email,. Chat boxes, etc.
- **Webchat Bots** - Appear on a website or social app that understands and responds to written text.
  - Limited scope of conversation and you can add friendly tone chit chat.
- **Telephone Voice Menus** - Speech Recognition and Speech Synthesis such as ordering from a restaurant or online banking.
- **Personal Digital Assistants** - Alexa, Siri, Etc.

Lesson 32: Conversational AI Services in Microsoft Azure

- **QnA Maker Service** - Automatically build a chat bot based on any semi-structured source.
  - Your website, product manuals, etc.
- **Azure Bot Service** - Enterprise-grade Chat Bots that will grow in sophistication and can be coded.