**AI-900 CheatSheet**

**Artificial Intelligence** a machine that performs jobs that mimic human behavior

**Machine Learning** a machine that gets better at a task without explicit programming

**Deep Learning** a machine that have an artificial neutral network inspired by the human brain to solve complex problems

**Data Scientist** a person with multi-disciplinary skills in math, statistics, predictive modeling and machine learning to make future predictions

**Dataset** is a **logical grouping of units of data** that are closely related and/or share the same data structure

* eg. MNIST, COCO

**Data Labeling** the process of **identifying raw data** (images, text files, videos, etc) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn

**Supervised Learning (SL)** Data that has been labeled for training

**Unsupervised Learning (SL)(uSL)** Data that has not been labeled, the ML model needs to do its own labeling

**Reinforced Learning (RL)** There is no data, there is an environment and an ML model generates data with many attempts to reach a goal

\*\*Neural Networks (NN) A network of nodes organized into layers (input, hidden, output) that is used to train ML models

**Deep Neural Network (DNN) A neural network that has \*\*3 or more hidden layers** is considered deep learning

**Backpropagation (BP)** Moves backward through the neural net adjusting weights to improve outcome on next iteration. How a NN learns

**Loss Function** A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

**Activation Functions** An algorithm applied to a hidden layer node that affects connected output e.g. ReLu

**Dense Layer** When the next layer increases the number of nodes

**Sparse Layer** When the next layer decreases the number of nodes

**General Processing Unit (GPU)** that is specially designed to quickly renger high-resolution images and video concurrently

* commonly used for non-graphical tasks such as machine learning and scientific computation

**Compute Unified Device Architecture (CUDA)** is a a parallel computing platform and API by NVIDIA that allows developers to use CUDA-enabled GPUs for general purpose computing on GPUs (GPGPU)

**ML Pipeline**

* Pre Processing - preparing data and feature engineering before passing data to an ML model for training or inference
  + Data cleaning - correcting errors within the dataset that could negatively impact the results
  + Data reduction - reducing the amount of data, or applying dimensionality reduction to reduce the dimensions of inputted vectors
  + Feature engineering - transforming data in numerical (vectors) to be ingested by an ML model
  + Sampling - balancing a dataset to be uniform across labels by adding or removing records
* Post Processing - translating the output of an ML model back into a human-readable format
* Training - the process of training the model
* Serving - the process of deploying a model to an endpoint to be used for inference
* Inference - invoking an ML model by sending a request and expecting back a prediction
  + Real-time endpoint - optimized for small or single item payloads, returns results quickly (a dedicated running server)
  + Batch-transform endpoint - optimized for larger batch predictions (server runs only for the duration of the batch)

**Forecasting**: Makes a future prediction with relevant data, analysis of trends, it's not "guessing"

**Predicting**: Makes a future prediction without relevant data, uses statistics to predict future outcomes, more of "guessing". Uses decision theory

**Performance/Evaluation Metrics** are used to evaluate different Machine Learning Algorithms

* Classification (Accuracy, F1 Score, Precision, Recall)
* Regression Metrics (MSE, RMSE MAE)

**Jupyter Notebook** A Web-based application for authoring documents that combine live-code, narrative text, equations, visualizations

**Classification** is a process of finding a function to **divide a labeled dataset into classes/categories**

* **Confusion matrix** is table to visualize the **model predictions** (predicted) vs **ground truth labels (actual)**

**Regression** is a process of finding a function to **correlate a labeled dataset into continuous variable/number**

**Clustering** is a process of **grouping unlabeled data based on similarities and differences**

**Cognitive Services** is an umbrella AI service that enables customers to **access multiple AI services** with an API key an an API Endpoint

**Decision**

* **Anomaly Detector** - Identify potential problems early on
* **Content Moderator** - Detect potential offensive or unwanted content
* **Personaliser** - Create rich, personalized experiences for every user

**Language**

* **Language Understanding** - Build natural language understanding into apps, bots, and IoT devices
* **QnA Maker** - Create a conversational question and answer layer over your data
* **Text Analytics** - Detect sentiment, key phrases and named entities
* **Translator** - Detect and translate more than 90 supported languages

**Speech**

* **Speech to Text** - Transcribe audible speech into readable, searchable text
* **Text to Speech** - Convert text to lifelike speech for more natural interfaces
* **Speech Translation** - Integrate real-time speech translation into your apps
* **Speaker Recognition** - Identify and verify the people speaking based on audio

**Vision**

* **Computer Vision** - Analyze content in images and video
* **Custom Vision** - Customize image recognition to fit your business needs
* **Face** - Detect and identify people and emotions in images

**Knowledge mining** is a discipline in AI that uses a **combination of intelligent services to quickly learn from vast amounts of information**

* **Ingest** content from a range of sources, using connectors to first and third-party data stores
* **Enrich** the content with AI capabilities that let you extract information, find patterns, and deepen understanding
* **Explore** the newly indexed data via search, bots, existing business applications, and data visualizations

\*\*Microsoft AI Principles (Responsible AI)

1. Fairness - AI systems should treat all people fairly
2. Reliability and Safety - AI systems should perform reliably and safely
3. Privacy and Security - AI systems should be secure and respect privacy
4. Inclusiveness - AI systems should empower everyone and engage people
5. Transparency - AI systems should be understandable
6. Accountability - People should be accountable for AI systems

**Common ML Workloads:**

* **Anomaly Detection** is the process of finding outliers within a dataset called an anomaly
* **Computer Vision** is when we use ML NN to gain high-level understanding from digital images or video
* \*\*Natural Language Processing (NLP) is ML that can understand the context of a corpus (a body of related text)
* **Conversational AI** is the technology that can participate in conversations with humans

**Azure Machine Learning Service** allows you to provision an ML studio to build and maintain ML models and pipelines

**Author**

* **Notebooks** - Jupyter Notebooks, an IDE to write Python code to build ML models
* **AutoML** - Completely automated process to build and train an ML model
* **Designer** - Visual drag and drop designer to construct end to end ML pipelines

**Assets**

* **Datasets** - data that you upload which will be used for training. Datasets can be versioned
  + **Open DataSets** are **publicly hosted datasets** that are commonly used for learning how to build ML models
* **Experiments** - Experiments are logical grouping of runs
  + **Runs** are ML tasks that perform on virtual machines or containers
* **Pipelines** - ML workflows you have built, or you have used in the Designer
  + Training Pipeline - pipelines to build and train an ML model
  + Inference Pipeline - pipelines that use a trained model to make predictions on real data
* **Models** - a model registry containing trained models that can be deployed
* **Endpoints** - when you deploy a model, it's hosted on an accessible endpoint eg. REST API
  + Real-time Endpoint - Invokes an ML model for inference
  + Pipeline Endpoint - Invoke the running on a Pipeline eg. CI/CD

**Manage**

* **Compute** - the underlying computing instances used for notebooks, training, inference
  1. **Compute Instances** - Development workstations that data scientists can use to work with data and models
  2. **Compute Clusters** - Scalable clusters of virtual machines for on-demand processing of experiment code
  3. **Inference Clusters** - Deployment targets for predictive services that use your trained models
  4. **Attached Compute** - Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters
* **Environments** - a reproducible Python environment for machine learning experiments

**Datastores** - **securely connect to your storage service** on Azure without putting your authentication credientials

* Azure Blob Storage, Azure File Share, Azure Data Lake Storage (Gen 2), Azure SQL database, Azure Postgres/MySQL database
* **Data Labeling** - have humans with ML-assisted labeling to label your data for supervised learning
  + Human-in-the-loop labeling
  + Machine-learning-assisted data labeling
* **Linked Services** - external services you can connect to the workspace eg. Azure Synapse Analytics

**Text Analytics**

* sentiment analysis find out what people think of your brand or topic
  + Labels include negative, positive, mixed or neutral
  + Confidence scores ranging from 0 to 1
* opinion mining granular information about the opinions related to aspects
  + granular data with a Subject and Opinion tied to a Sentient
* key phrase extraction quickly identify the main concepts in text
* language detection detect the language an input text is written in
* named entity recognition (NER) - detects **words and phrases mentioned in unstructured text** that can be associated with one or more semantic types

**Language Understanding (LUIS)** is **a no-code ML service to build natural language into apps**, bots, and IoT devices

* Natural Language Understanding (NLU) - the ability to *transform* a linguistic statement into a representation that enables you to understand your users naturally
* LUIS key schema components
* **intentions** - what the user is asking for
  + a LUIS app always contains a None Intent
* **entities** - what parts of the intent is used to determine the answer
* **utterances** - examples of user input that includes intent and entities to train the ML model to match predictions against real user input

**QnA Maker** generate a bot from a URL, PDF, or DOX DOCX

* multi-turn conversation - follow up prompts to narrow to a specific answer
* Chit-chat - personalized canned responses

**Azure Bot Service** - allows you to host bots

* Bot Framework SDK - an end-to-end SDK to design, build, test, publish, connect and evaluate bots
* Box Bot Framework Composer - a desktop application to design bots, leverages the Box Bot Framework SDK