Maxine Hickey
October 2, 2023
Week 6: Introduction to Large Language Models
What are Large Language Models (LLMs)?
- Deep learning algorithms that can recognize, summarize, translate, predict and generate text and other forms of content based on knowledge gained from massive datasets
- One of the most successful transformer models

What are LLMs used for?
- Human communication!!
- They broaden AI's reach across industries
- They can help generate complex solutions for the world's toughest problem

How do LLMs work?
- Learn from large datasets (but large is relative)
- All of this text/data is fed into the algorithm using unsupervised learning
- Unsupervised Learning: a dataset without explicit instructions on how to handle it (unlabeled data)
    - Less relevant to this and more to data mining and general ML: supervised learning and reinforcement learning
        - Supervised Learning: trains a model on labeled data
            - Used in image classification, regression, and object detection
        - Reinforcement Learning: trains a model to make decisions by interacting with an environment
            - Used in game playing, robotics, and autonomous vehicles

Most Used Apps for LLM
- [ChatGPT](#)
- [Gemini](#)
- [Perplexity AI](#)
- [NVidia Article](#)

Some challenges of LLMs
- Scaling and maintaining can be difficult and expensive
- Find large enough datasets to be able to fully train the LLM
- NVidia GPU sometimes doesn't work :/

[Text Prompting Intro Code](#)
- Colab notebook but will also have a .py file