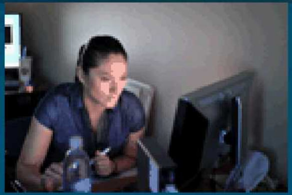
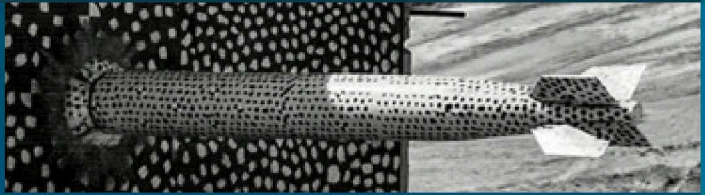
SAND2020-7359PE

A Summer in Machine Learning





Presented By

Sean Timm

Overview

- Mathematics Course with Connor Frost & Patrick Cooper.
  - Helped to create a course on the mathematics behind machine learning. This involved filming Khan Academy style videos, creating detailed lecture notes, and creating homework problems.

- Learning Machine Learning and Implementing Parts of an ML Library.
  - Library seeks to abstract away some of the details of writing machine learning models in TensorFlow. A plethora of the code written ends up being rewritten every time you create a model – this library seeks to allow for more code reuse and less duplicate code.

The Mathematics Course

Overview

The Joys of Math

Often when teaching mathematics, the purpose is unclearly defined. Mathematics rarely has to justify the purpose of learning it (which, in academic situations, is often left as a proof to the reader). The course we produced is different, in that it had a clearly defined goal:

"Provide mathematical justification for all structures which allow for the operation of machine learning."

Which creates the demand for more effort to be focused upon the tangible and applications of the subject area, in direct correspondence with our goal. This creates the need to adjust what subject matter you cover, so as to produce content that is both meaningful for the viewer while also being efficient.

Course Overview

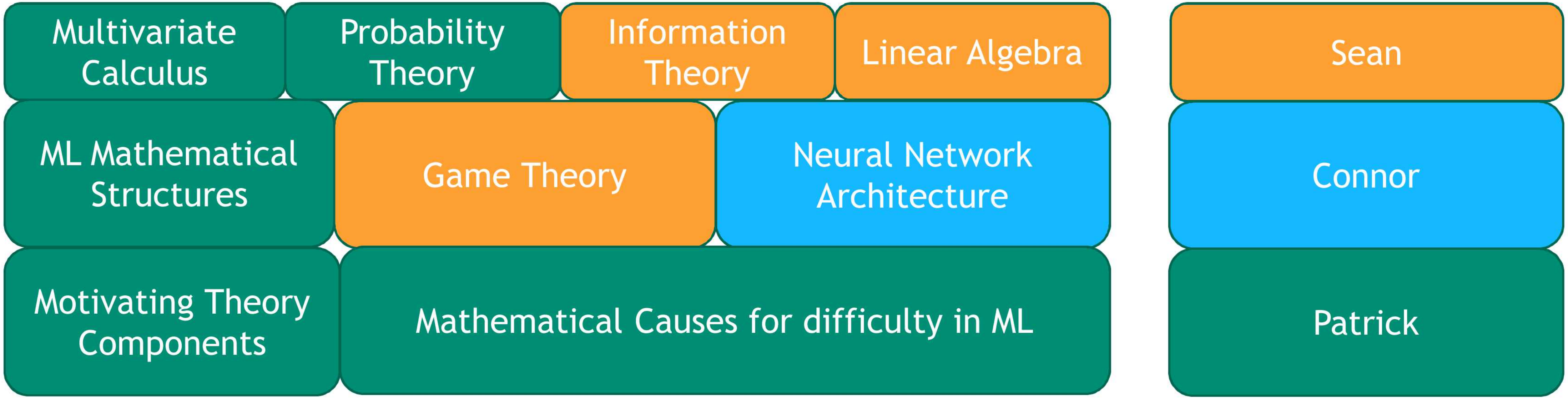
Mathematically, machine learning is diverse in what subject matter pertains to it and facilitates its operations. We split it up into sections and each of us took sections to cover.

- Multivariate Calculus
- Probability Theory
- Information Theory
- Linear Algebra
- ML Mathematical Structures
- Game Theory
- Neural Network Architecture
- Motivating Theory Components
- Mathematical Causes for difficulty in ML

*Size does not represent time or effort spent.

Course Overview

Mathematically, machine learning is diverse in what subject matter pertains to it and facilitates its operations. We split it up into sections and each of us took sections to cover.



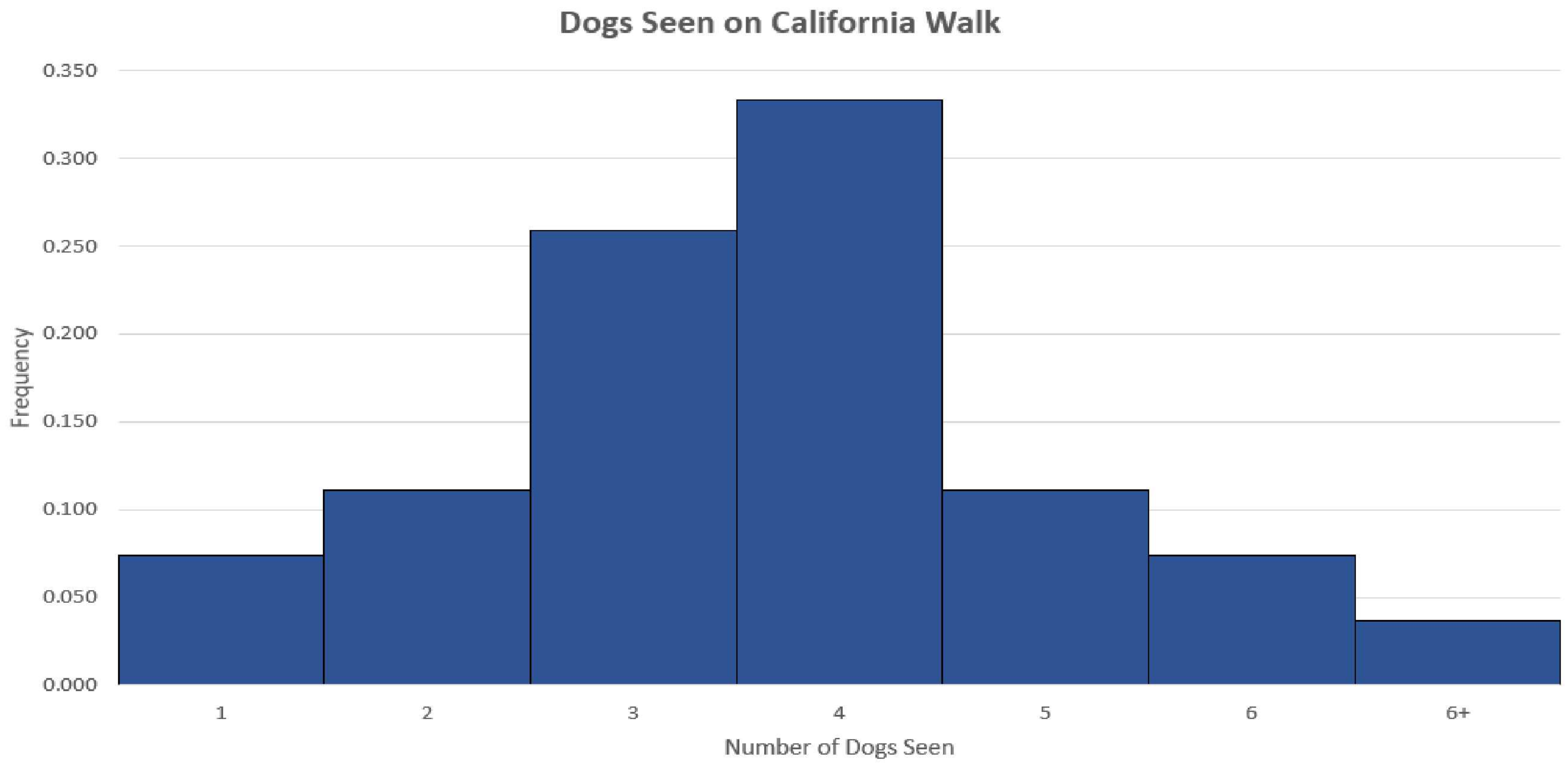
*Size does not represent time or effort spent.

Information Theory - Example

Kullback-Leibler Divergence

A Specific Subject: Kullback-Leibler (KL) Divergence

Before I begin getting into KL Divergence, here is an example of a probability distribution:



A Specific Subject: Kullback-Leibler (KL) Divergence

What KL Divergence does; it gives us a measure of how well one probability distribution can represent another. The formula for discrete distributions:

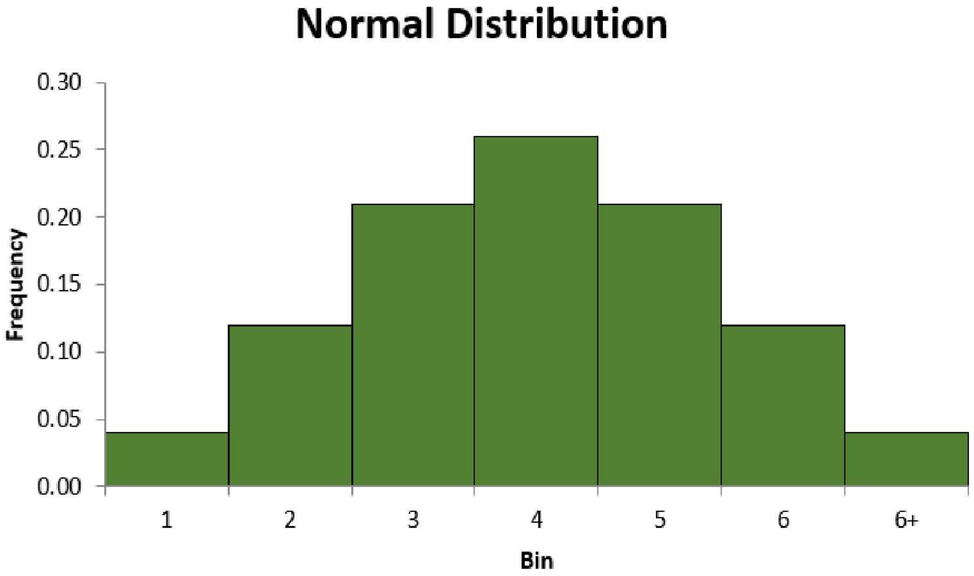
$$D_{KL}(p||q) = \sum_{i=1}^{N} p(x_i) \cdot log \frac{p(x_i)}{q(x_i)}$$

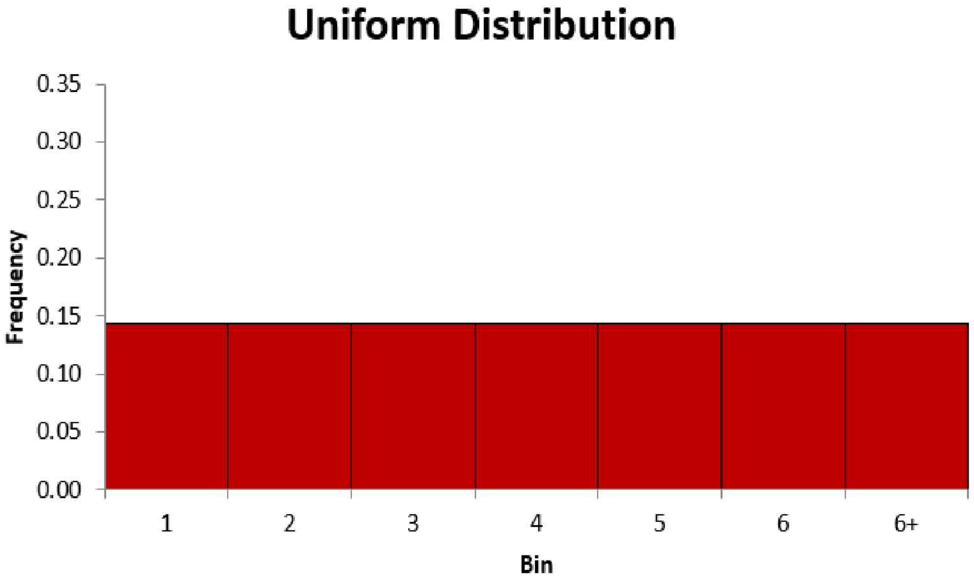
p(x) = Probability of x occurring in the probability distribution p.

q(x) = Probability of x occurring in the probability distribution q.

Intuition of the formula: sum of the logarithmic differences, with each scaled by how frequent the particular x occurs in p. The lower the returned measurement, the better the representation.

A Specific Subject: Kullback-Leibler (KL) Divergence





So, if we wanted to choose a distribution to represent our dog distribution, which of these would be preferable to KL Divergence?

Running the formula, we get:

- 0.028~ for the normal distribution.
- 0.101~ for the uniform distribution

We can see that the normal distribution is significantly lower than the uniform distribution and provides a much better representation.

How it applies to Machine Learning

KL Divergence has a multitude of applications in Machine Learning:

- Supervised Learning.
- Reinforcement Learning.
- Information Bottlenecks.
- Variational Auto-Encoders (VAE).
- Generative Adversarial Networks (GAN).
- Synthetic Data Generation.

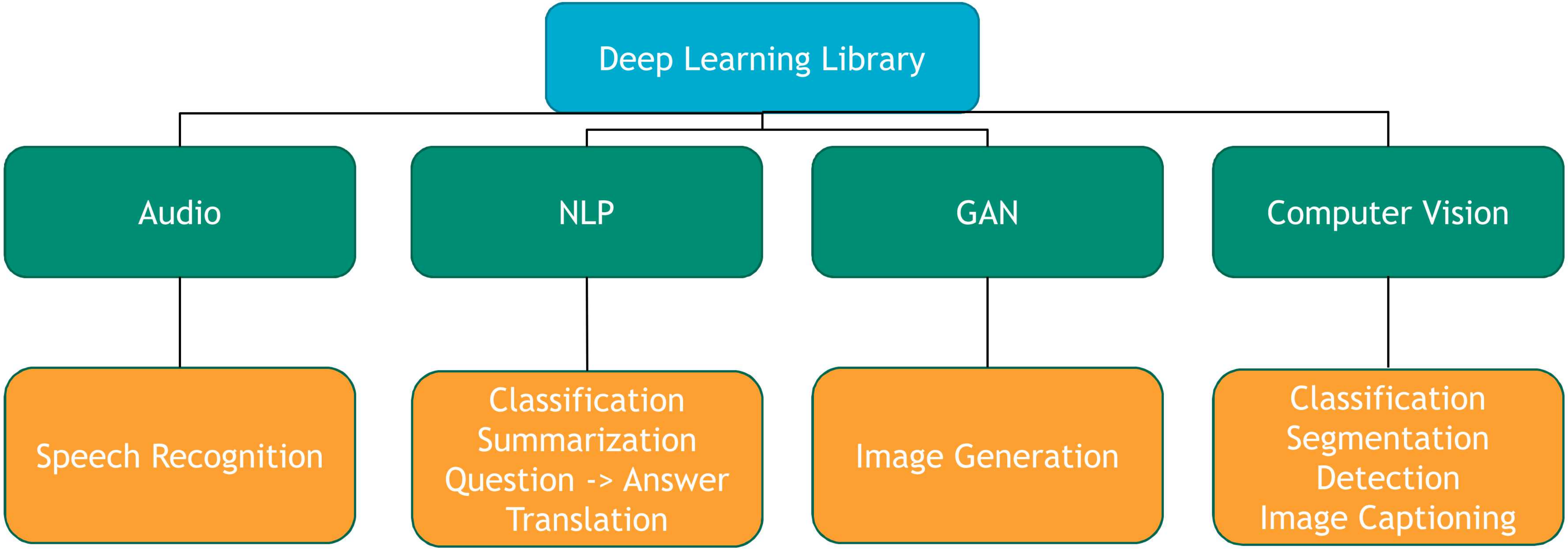
Deep Learning Library

Working Towards Abstraction

Library Overview

Library development consisted of designing modules to facilitate machine learning's four common use cases:

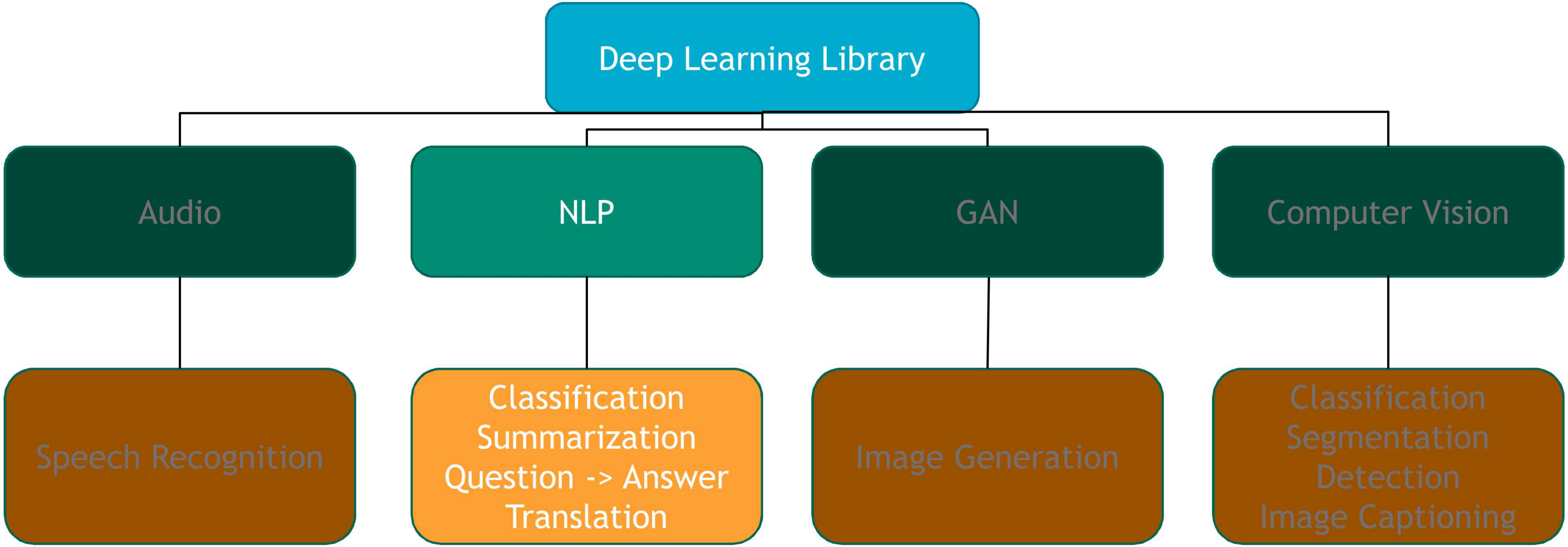
- Computer Vision
- Natural Language Processing (NLP)
- Audio Processing
- Generative Adversarial Networks (GANs).



Library Overview

Library development consisted of designing modules to facilitate machine learning's 4 common use cases:

- Computer Vision
- Natural Language Processing (NLP)
- Audio Processing
- Generative Adversarial Networks (GANs).



Classification

My primary focus for the early stages of the summer was text classification. The goal of text classification is; given a passage of text, classify it based on some labels. A few example scenarios:

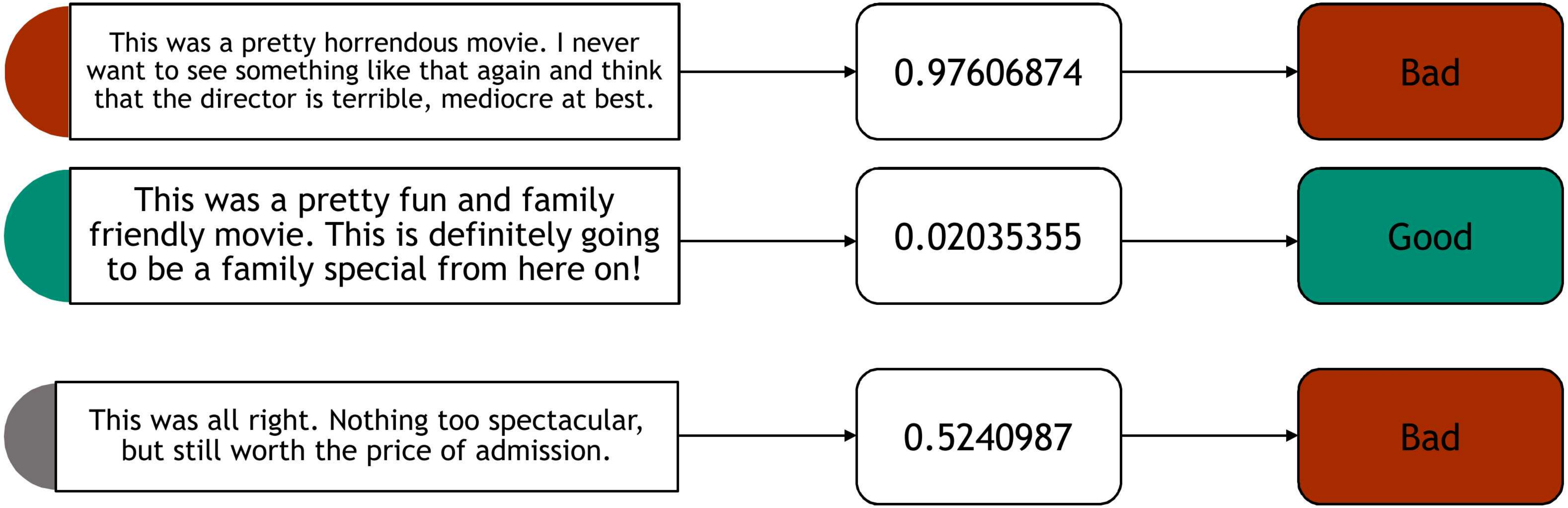
- Given a sentence, determine if it is sarcastic.
- Given a review, determine if it is positive or negative.
- Given some code, determine what language it was written in.

Text classification is frequently used to analyze the sentiment of text, with that being its most common use case.

One of my biggest challenges was not only learning ML principals and how to actually write the code to apply it, but to also envision abstractions that would be beneficial for our library.

Concrete Example

If we were in-person, I would've loved to have been able to let the user enter in phrases and let the model categorize them live, but this will have to suffice. Here is some example output from a model that is classifying movie reviews as good or bad. Bad reviews are closer to 1, good reviews are closer to 0.



Notice how the review that was somewhere in between good and bad was near 0.5? However, it is still closer to 1 than 0, so it was marked as a bad review.

Questions?