

Tech Feasibility Draft

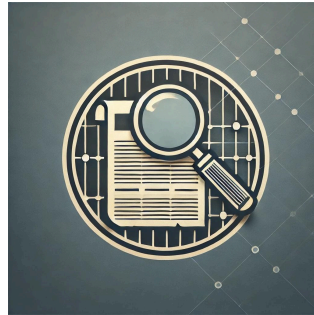
Team: INSIGHT

Sponsor: Mike Taylor

Faculty Mentor: Scott Larocca

Team Members:

Joshua VanderMeer, Michael Vertin,
Aidan Hebert, Forrest Hartley



Date: 10/25/2024

Table of Contents:

Introduction: (.75-1.5 pages)

<Big Picture>

<How the problem is being dealt with>

<General description of our product/approach>

Transition into the main document body

<Talking about project in general> -> <What role the document plays>

<Outline flow of document>

Technological Challenges:

<list of high-level requirements/hurdles>

Technology Analysis: (3-5 pages)

<Major technological issues/design decisions>

Introduce the issue:

Desired Characteristics:

Alternatives:

Analysis:

Chosen Approach:

Proving Feasibility:

Technological Integration:

<Bring it all together>

<How will all of these micro-solutions come together>

<System Diagram> - how major elements relate

Conclusion:

<Importance of problem>

<Summary of document>

<Project steps>

Vector Database Management:

- [ScaNN](#)
- [Pinecone](#)
- [Faiss](#)
- [Milvus](#)
- [pytensor](#)
- [scikit](#)

Convolutional Neural Network (CNN):

- Only images
- Detects features of images
- The embedding represents how well images pair with specific terms
- All input have the same length
- Maximizes accuracy with a 10M image training set
- [TensorFlow](#)

Transformers:

- Can use different data types
- The embedding represents the meaning of the media
- Quicker to convert to an embedding
- Continues to improve with over 10M images
- [Hugging Face](#)
- [TensorFlow](#)

Pretrained Models:

- Stores results of training on generic data sets
- Can be fine-tuned to smaller more specific data sets
- [Hugging Face](#)

ContentDM

DigitalOcean

AWS