WGU C951

Task 3

MACHINE LEARNING PROJECT PROPOSAL

Name: Cameron Woodard

Student ID # : 003443434

Date: 08/15/23

Table of Contents

A. Project Overview....................................................................................................................pg#2 A.1. Organizational Need ......................................................................................................... pg#2 A.2. Context and Background.................................................................................................... pg#2 A.3. Outside Works Review ...................................................................................................... pg#2 A.4. Solution Summary............................................................................................................. pg#3 A.5. Machine Learning Benefits................................................................................................ pg#3 B. Machine Learning Project Design......................................................................................... pg#3 B.1. Scope ................................................................................................................................. pg#3 B.2. Goals, Objectives, and Deliverables ................................................................................. pg#4 B.3. Standard Methodology ..................................................................................................... pg#5 B.4. Projected Timeline ............................................................................................................. pg#5 B.5. Resources and Costs........................................................................................................... pg#6 B.6. Evaluation Criteria ............................................................................................................ pg#7 C. Machine Learning Solution Design....................................................................................... pg#8 C.1. Hypothesis.......................................................................................................................... pg#8 C.2. Selected Algorithm............................................................................................................. pg#8 C.2.a Algorithm Justification..................................................................................................... pg#8 C.2.a.i. Algorithm Advantage ................................................................................................... pg#8 C.2.a.ii. Algorithm Limitation .................................................................................................. pg#9 C.3. Tools and Environment ..................................................................................................... pg#9 C.4. Performance Measurement................................................................................................. pg#9 D. Description of Data Sets ..................................................................................................... pg#9 D.1. Data Source ..................................................................................................................... pg#9 D.2. Data Collection Method................................................................................................... pg#9 D.2.a.i. Data Collection Method Advantage............................................................................ pg#10 D.2.a.ii. Data Collection Method Limitation .......................................................................... pg#10 D.3. Data Normalization.......................................................................................................... pg#10 D.4. Data Security.................................................................................................................... pg#10 References................................................................................................................................ pg#11

**A. Project Overview**

**A.1. Organizational Need**

My project will be implementing an image recognition AI for a stock photo company using machine learning to identify objects in the photos. The company manages a large collection of photos, and they want to enhance the platform’s searching and categorization capabilities. They would like the process to become automated to improve the experience for customers and employees alike.

**A.2 Context and Background**

The company is experiencing an influx of stock photos that need to be organized and tagged to meet the customer demand. The current way to categorize photos is manual and does not keep up with the demand. The company’s labor resources can’t keep up, leaving a backlog of stock photos that can’t make it to the company’s online marketplace. The context of this project is to advance the company’s technology and satisfy the need for categorizing the stock photos. Through the use of image recognition AI, the company wishes to upload more photos and have a larger variety than its competitors.

**A.3. Outside Works Review**

In order to better research the development of image recognition AI, three outside works related to machine learning and object recognition were viewed. They provided insight to help achieve efficient image categorization.

1. “A Gentle introduction to object recognition with deep learning” by Jason Borwnlee, machinelearningmastery.com (1)

This article provides an introduction to the different various methods of computer vision used for AI image recognition. It goes into deeper defining object recognition, region-based convolutional neural networks (R-CNNs), and You Only Look Once (YOLO) techniques for object recognition. The article’s various approaches to image recognition provides an understanding of how image recognition can be used to complete the project’s tasks.

2. “Faster R-CNN: Towards Real-Time-Object Detection with Region Proposal Networks” by Shaoquing Ren, Kaiming He, Ross Girshick, and Jian Sun (2)

This paper goes into Faster R-CNN, which enables accurate object detection in images, which is needed for tagging and categorizing photos for the stock photo company. It is superior in speed to regular R-CNN, boosting efficiency in an environment where a large number of photos are going to need to be processed quickly. The real-time capability of Faster R-CNN coincides with the stock photo company’s need to quickly categorize the photos for the search function of the company’s website. The methodology in the paper will help guide the team to implement a quick and effective image recognition system.

3. “You Only Look Once: Unified, Real-Time Object Detection” by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi (3)

You Only Look Once (YOLO) can process images in real-time at 45 frames per second, learning very general representation of objects. This makes it a good contender for automatically tagging and categorizing images, the ultimate goal of the machine learning project to categorize photos for the company. While YOLO’s speed is impressive, the paper is used to note the limitations of YOLO and provide context on why choosing Faster R-CNN for image recognition is the better option. Considering YOLO contributes to the decision-making process of the project and helps achieve a recognition solution that is both quick and accurate.

Exploring R-CNN, Faster R-CNN, and YOLO provide the company a better understanding of the options available to them for the image recognition AI. The evaluation of the benefits and drawbacks will help the company come to a better decision on which to use for the project.

**A.4. Solution Summary**

Due to its higher level of accuracy and object detection, Faster R-CNN would be the solution for the image recognition software. While YOLO works faster, it is not as accurate in identifying objects in a scene as Faster R-CNN. This decision also depends on the resources of the company: Faster R-CNN will be more manageable in terms of scale, but requires more resources. If the company can afford it, then faster R-CNN will be better due to its accuracy. Faster R-CNN uses a two-stage approach in identifying objects, creating region proposals and classifying objects within those proposals.

**A.5. Machine Learning Benefits**

While Faster R-CNN training is expensive where space and time are concerned and object detection is slow, there are several benefits in using this machine learning. It uses a region proposal network to bring attention to the Fast R-CNN network, informing it where to pay attention in the image. Faster R-CNN can be adaptable to various images and objects, allowing it to handle a wide range of tags used by the company This makes it a good choice for fast and accurate detection of objects in images for categorization by the stock photo company.

**B. Machine Learning Project Design**

**B.1. Scope**

The scope of the project will be implementing the image recognition system to allow enhanced searching and categorization of photos for the stock photo company.

In scope items include:

* Collecting data and preparing the stock photos to be used for AI training
* Implementing the Faster R-CNN for object detection and image categorization
* Integrating the AI with the website to allow users to search for photos

Out of scope items include:

* Having the AI detect the contents of text within photos. The content of text in images does not pertain to how the photo is categorized for the company.

**B.2. Goals, Objectives, and Deliverables**

The primary goal of the project is to enhance the stock photo company’s categorization and searching functions by implementing an image recognition AI using machine learning. This will lead to optimized search results, and better overall operational efficiency. Here are some of the objectives to be completed, and the deliverables and goals related to them:

Objective 1) Develop an accurate image recognition System

Deliverables from this objective include the implementation of Faster R-CNN image recognition, and a trained model to accurately identify objects within the stock photos. This contributes to the overarching goal of increasing the efficiency and accuracy of the stock photo management program. Another goal by implementing the AI includes increasing customer satisfaction with the program, as the search results will be more in line with what the customer was looking for.

Objective 2) Automating tagging and categorization

To tag and categorize the photos, an API deliverable will be created. The integration of the image recognition AI with the company’s website will also have been created. These two deliverables will contribute to the goal of making the company’s website more state-of-the-art, and gaining a foothold over the competition.

Objective 3) Improving the search functionality

The deliverable from this objective is the enhancement of the website’s search function. This will contribute to the ease-of-use of the website and helps achieve the goal of increasing customer satisfaction.

Objective 4) Reducing labor costs

By implementing the image recognition AI, we will be reducing the manual tagging efforts. We will also be able to eliminate the backlog of photos needing to be tagged in the company’s collection.

Other deliverables include the documentation for the image recognition model, including any training which will be needed for the employees. There should be enough information to allow employees to manage the image recognition model, tweaking it as the business needs. All of these deliverables and objectives will help achieve the company’s goals: improving customer satisfaction, reducing labor costs, and increasing profits for shareholders.

**B.3. Standard Methodology**

SEMMA is often used in machine learning projects to improve performance and deliver more useful services to customers (4). To apply this to my project, here are the different stages of SEMMA in context of the image recognition project:

**Sample -** A sample of the dataset will be given in order to construct the machine learning model. In this case, a set of stock photo images will be used for training and validation.

**Explore -** In this step, analysis is done to study relationships between data elements and identify gaps in the data. A closer look will be taken and the trends and patterns of the data that might have an effect on the model’s accuracy.

**Modify -** Here, the data is parsed and cleaned to be passed onto the modeling stage. We also check if it needs any refinement or transformation. This might include augmenting the dataset by adjusting the images themselves to increase the variation of the dataset.

**Model -** Modeling applies data mining techniques to produce a projected model of how data achieves the desired outcome of the process. For this part, we will select an image recognition model architecture, and train it using the prepared dataset.

**Assess -** Finally, we evaluate the model in its reliability. The performance of the model will be compared against our project’s objective (tagging images based on their contents).

**B.4. Projected Timeline**

The start and end dates and length of the project will depend on the size of the team, the resources to be used, and any unforeseen challenges. However, we can make a rough estimate for each of the different tasks to be completed:

Task 1) Preparing and collecting data (3 weeks) [Start 01/01/24 End 01/22/24]

* Define the projects scope and objectives
* Gather images to be used for the dataset
* Format the data for project use

Task 2) Develop the image recognition model (5 weeks) [Start 01/23/24 End 02/27/24]

* Research and select the Faster R-CNN Model to be used
* Set up the development environment and tools
* Train the Faster R-CNN Model

Task 3) Integration and Testing (3 weeks) [Start 02/28/24 End 03/20/24]

* Develop the API used for the automating tagging of photos
* Test the image recognition model for accuracy in object recognition
* Obtain initial testing with users and gain feedback

Task 4) Optimization and Refinement (2 weeks) [Start 03/21/24 End 04/04/24]

* Fine-tune the model to improve performance
* Enhance the accuracy of tagging by addressing user feedback
* Optimize the model’s inference speed

Task 5) Deployment and Evaluation (1 week) [Start 04/04/24 End 04/11/24]

* Deploy the model into the production environment
* Monitor the system’s performance
* Collect feedback from the users about the search functionality

Task 6) Reporting and wrap-up (1 week) [Start 04/12/24 End 04/19/24]

* Create a summary of achievements and outcomes
* Create the project documentation
* Present the results and the project’s completion

**B.5. Resources and Costs**

| **Resource** | **Description** | **Cost** |
| --- | --- | --- |
| GPU server | Used to train the image recognition model using Faster R-CNN | $8,500 (4) |
| Development tools | An IDE is required to develop the programs used. These could be free or paid depending on the IDE used. | Free - $500 |
| Labor | This includes paying software developers and team members which research, develop, test, and manage. This amount varies by how long the project lasts. | > $2,000 |
| Cloud services | Depending on how many stock photos the company manages, it would make sense to pay a monthly fee to have the data stored in the cloud by a third party. | $500 monthly |
|  | **Total** | > $11,500 |

**B.6. Evaluation Criteria**

The following key performance indicators are used to evaluate the effectiveness of the project:

| **Objective** | **Success Criteria** |
| --- | --- |
| Image recognition accuracy | This is the primary objective of the project, enhancing the accuracy of stock photo image categorization. The success criterion for this will be an accuracy rate of 90% in identifying and tagging objects within the photos. |
| Labor savings | With the automation of image tagging, employees will have more time to work on other tasks. This will reduce the amount of overhead and labor required to tag images manually. The success criterion for this will be a reduction of at least 60% in labor hours dedicated to manually tagging photos. |
| Increased search result accuracy | The accuracy and efficiency of tagging will lead to better search results when browsing stock photos, ensuring customers are more able to find the image results they’re looking to purchase. The success criterion will be a 30% increase in the relevancy of search results. This is reported by user feedback and compared against previous search performance. |
| Return on investment | A major part of the project’s success will depend on its impact on the company’s finances. The successful criteria for the return on investment will be a calculated cost reduction of at least 25% in labor expenses related to photo tagging and categorizing in the first year the AI is implemented. |

**C. Machine Learning Solution Design**

**C.1. Hypothesis**

By implementing Faster R-CNN and an image recognition AI using machine learning, the stock photo company will improve its photo categorization and searching capabilities. This will reduce the amount of manual tagging necessary, and enhance the user experience. The testing that will demonstrate this hypothesis includes user experience testing and scalability testing. The user experience testing includes gaining feedback in the accuracy of search results, as well as the program’s ability to correctly identify and categorize stock photos. The scalability testing will be done by providing a large set of photos for the algorithm to identify and categorize.

**C.2. Selected Algorithm**

Faster R-CNN, like that used in the project, is a type of supervised learning. Supervised learning uses labeled datasets to train algorithms to classify data or predict outcomes (6). In this context, a dataset of images are inputted for the algorithm to learn to identify what is in those images. This is different from unsupervised learning, which uses unlabeled data. (6)

**C.2.a Algorithm Justification**

The advantage of using Faster R-CNN is that it uses a region proposal network to bring attention to the Fast R-CNN network, informing it where to pay attention in the image. Faster R-CNN can be adaptable to various images and objects, allowing it to handle a wide range of tags used by the company This makes it a good choice for fast and accurate detection of objects in images for categorization by the stock photo company.

**C.2.a.i. Algorithm Advantage**

The degree of confidence in the algorithm will be expressed by having employees analyze the correctness of the image recognition AI using their expert judgment.

**C.2.a.ii. Algorithm Limitation**

Faster R-CNN is computationally more expensive than other image recognition options. YOLO recognition could be faster, but would be at the cost of the AI’s ability to recognize objects. The accuracy of Faster R-CNN is why it was chosen over other options.

**C.3. Tools and Environment**

The operating system used to develop the AI will depend on the company’s preferences. Popular choices include Linux, MacOS, and Windows. The programming language used can also vary, but Python is often used for machine learning and would be an excellent choice. Any sort of license to use Faster R-CNN will need to be acquired in order to start training the image recognition software.

**C.4. Performance Measurement**

To measure the performance of the machine learning solution, the AI’s accuracy will be assessed. It should be able to correctly identify and categorize the photos, with as few errors as possible. The program will need to be continuously monitored to look for areas of improvement and ways to make the algorithm more accurate in its predictions and categorizations.

**D. Description of Data Sets**

**D.1. Data Source**

The sources of the data to be used for the image recognition project include the in-house photo collection owned by the stock photo company, and outside sourced image datasets available online. These will be used to train the AI to detect objects within photos and categorize them.

**D.2. Data Collection Method**

To collect the data, we will need to acquire previously tagged images to train the model, and validate that the image recognition component is working correctly. We will be training the Faster R-CNN to work with real-world examples, ensuring it is ready when it is deployed into the live environment.

**D.2.a.i. Data Collection Method Advantage**

Using the stock photo company’s existing photo library to train the model will make it adjusted to the specific requirements that the company is looking for. If the photos are exclusively of nature, for example, it would not make sense to train the model to identify the model of cars in photos, as that's not the type of photos the stock photo company collects. Training to work with the type of photos the company deals with will make it more able to handle what it will be working with on a day-to-day basis.

**D.2.a.ii. Data Collection Method Limitation**

The limitation of using the in-house dataset is that it might be limiting the object recognition model’s ability to handle outliers and items it might not be familiar with. For example, if the items within an image do not correlate to any tags used by the stock photo company, the AI might not understand what it is looking at and it will require human intervention to manually tag the photo or remove it if it is a type of photo the company chooses not to offer in it’s business. We need to make sure that the model is able to identify and categorize a variety of photos, so customers will be better able to search for them when they are correctly categorized.

**D.3. Quality and Completeness of Data**

We need to prepare the data before it is trained into the model. This includes formatting the dataset into a way the image recognition model can accurately understand, and reduce the computational time needed to look at each photo of the dataset. Any outlier stock photos and edge cases need to be carefully watched over to ensure they are correctly categorized, as these are representative of the worst-cases the image recognition AI might face in the live environment. The AI must have access to all the tags the stock photo company uses to categorize their photos, as limiting the amount of categories for the photos will result in less accurate search results for the customers later on.

**D.4. Precautions for Sensitive Data**

When transferring the stock photos the company sells across employees, it is imperative that the data be encrypted so that none of the stock photos get stolen while being transferred. We must also limit the access to the stock photos to only those who need it to complete the project; these photos are the livelihood of the business and must be kept safe at all times. When communicating on the project, employees must go through the approved channels and avoid working on their own personal devices. Doing so could risk the security of the business, exposing the business’s secrets to its competitors or those who wish to commit a cyberattack or steal sensitive data.

**References**

1. Brownlee, Jason. “A Gentle Introduction to Object Recognition with Deep Learning.” *MachineLearningMastery.Com*, 26 Jan. 2021, machinelearningmastery.com/object-recognition-with-deep-learning/.
2. Ren, Shaoqing, et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” *arXiv.Org*, 6 Jan. 2016, arxiv.org/abs/1506.01497.
3. Redmon, Joseph, et al. “You Only Look Once: Unified, Real-Time Object Detection.” *arXiv.Org*, 9 May 2016, arxiv.org/abs/1506.02640.
4. Hotz, Nick. “What Is SEMMA?” *Data Science Process Alliance*, 12 Aug. 2023, www.datascience-pm.com/semma/.
5. Boesen, Michael Reibel. “GPU Servers for Machine Learning Startups: Cloud Vs on-Premise?” *Medium*, 14 Jan. 2019, medium.com/@thereibel/gpu-servers-for-machine-learning-startups-cloud-vs-on-premise-9a9dedfcadc9.
6. “What Is Supervised Learning?” *IBM*, www.ibm.com/topics/supervised-learning. Accessed 16 Aug. 2023.