Project Proposal

Nowday's e-commerce Marketplaces make our life easier and give us more options with various prices. Moreover, it is being part of our daily life.

The problem

These Marketplaces are facing many difficulties on both sides (user experience/marketplace company), which may effect in its journey experience. Some of these difficulties:

- **Distraction and duplication in a product category.**

- **Monitoring and Avoiding prohibited products.**

1. Distraction and duplication in product category

For users: If the user trying to create a new sell, the user will go through sell form that includes (Images, Title, Description, Category, etc.). In the category field, it may ask to type text or select from a dropdown list. All of these options can be caused a problem.

Type text	Select from list

1. Typo mistakes:
    e.g. Shoes -> (shoos, shows) !
2. Out of categories:
    e.g. Shoes -> (Men's Nike Shoes Air Force LV8) !
	1. Get Confusion:
	    e.g. Should I chose brand or shoes or men section?

For marketplace platform:

| Type text	Select from list
1. Expand database when each user insert a new category. 2. Make duplicated categories.	1. Loss a lot of products categorization when user choose wrong category.

2. Monitoring and Avoiding prohibited a products

Restricted products are always a sensitive topic for users and platform. Some examples of the prohibited items (Drugs, Guns, Trafficking in Human Beings, etc.).

With the growth of products on the platform, it will be more difficult to monitor or judge for prohibited items manually.

Project solution

In this project, I am trying to achieve a good solution for previous difficulties by using Machine Learning (ML) with some tools that can provide Image Classification suggestions for product categories, also Image Classification and Text Classification for monitoring and avoiding prohibited products quickly with automatic detection.
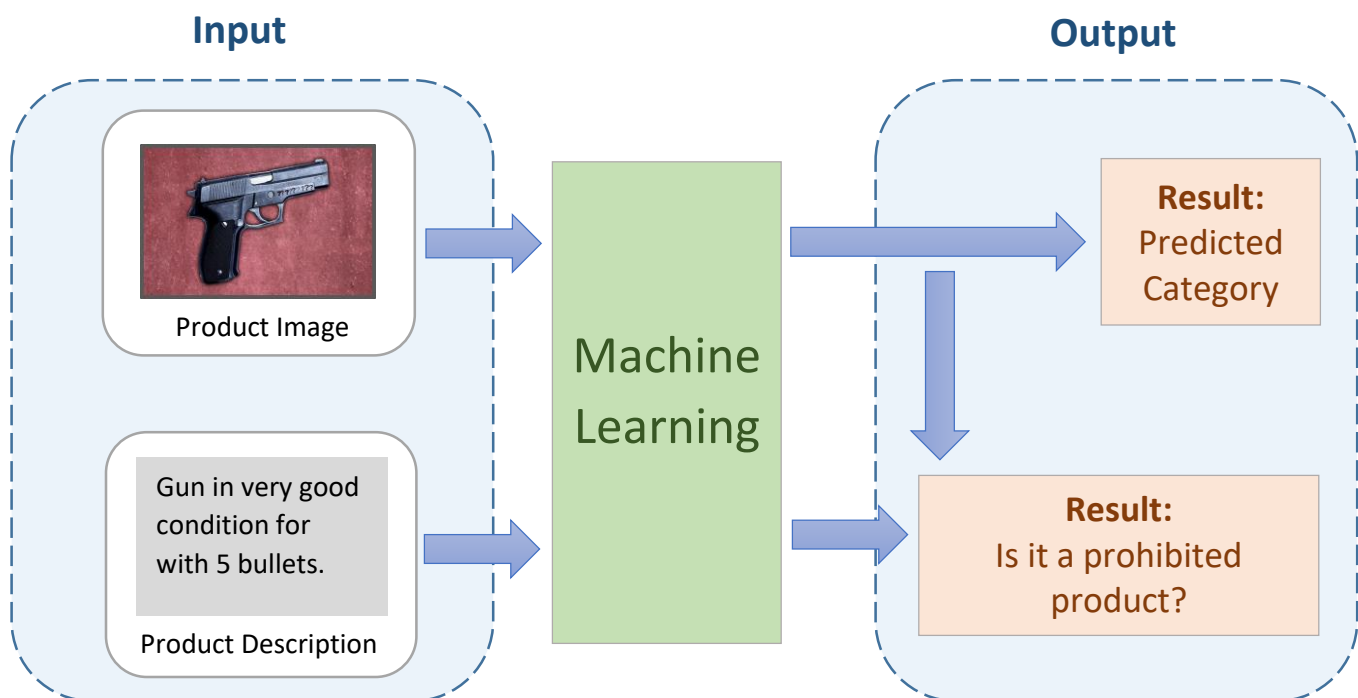
Datasets

For image classification, I am going to use:

1. **CIFAR-10 dataset:** It contains of **60000** (32x32) color images in 10 classes. and it will be for first stage. **link**

2. **ImageNet dataset:** It is a common dataset contains over **1 million** images. **link**

For text classification, I am going to create my own dataset.

Tools

I will conduct the experiment by using:

- **Environment:** Jupyter Notebook.

- **Programming Language:** Python.

- **Libraries:** Pandas, NumPy, Sklearn, Matplotlib, Tensorflow.

- **Other:** Convolutional Neural Network (CNN), Multiplayer Perceptron (MLP).

MVP

- Prepare the environment
- Import all libraries and dependencies
- Load CIFAR-10 dataset
- Build a CNN
- Repeat previous steps with ImageNet dataset
- Apply the inception-v3
- Apply MLP
- Measure scores and results