INSY 670 | Final Group Project

McGill | DESAUTELS

Maximizing Digital Engagement & Conversions

A Data-Driven Approach to Enhancing Ad Effectiveness on Instagram

Social Media Analytics | Section 075
Winter 2024

Presented to: Prof. Taha Havakhor

Adrian Alarcon Delgado
Jiaxuan Wang
Meriem Mehri
Siqi Wang
Xiaorong Tian
Zhicheng Zhong

McGill | DESAUTELS

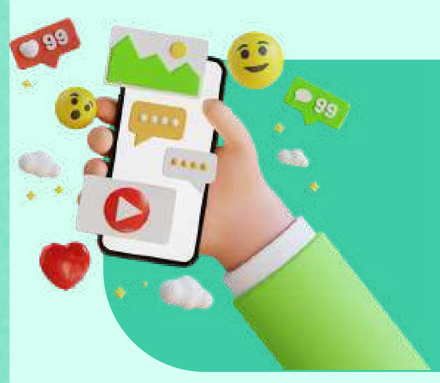
Agenda

- Project Overview & Problem Statement
- Data Collection & Curation Process
- Text & Image Modeling - SVD & CNNs — *Demo*
- Results & Insights Interpretation
- Business Outcomes & Implications

McGill | DESAUTELS

Executive Summary

Our project focuses on optimizing social media ads through data-driven strategies. By analyzing the interplay between text and image elements, we aim to boost engagement and conversions. Leveraging a diverse dataset from Instagram, we employ advanced analytics to extract actionable insights for ad content creation. Our findings offer businesses precise guidelines to enhance ROI and brand visibility in the competitive digital landscape.



Project Context

Utilize data analytics to optimize social media ad content for increased engagement and conversions, analyzing the interaction between text and image elements in Instagram ads.

- **Approach:** Develop a robust analytical framework to systematically assess content effectiveness and identify patterns for more impactful advertising.
- **End Goal:** Provide actionable insights to guide influencers and businesses in crafting compelling ad content that drives desired actions.

Problem Statement

Businesses struggle to create ad content on Instagram that resonates with audiences and drives conversions, crucial for brand growth.

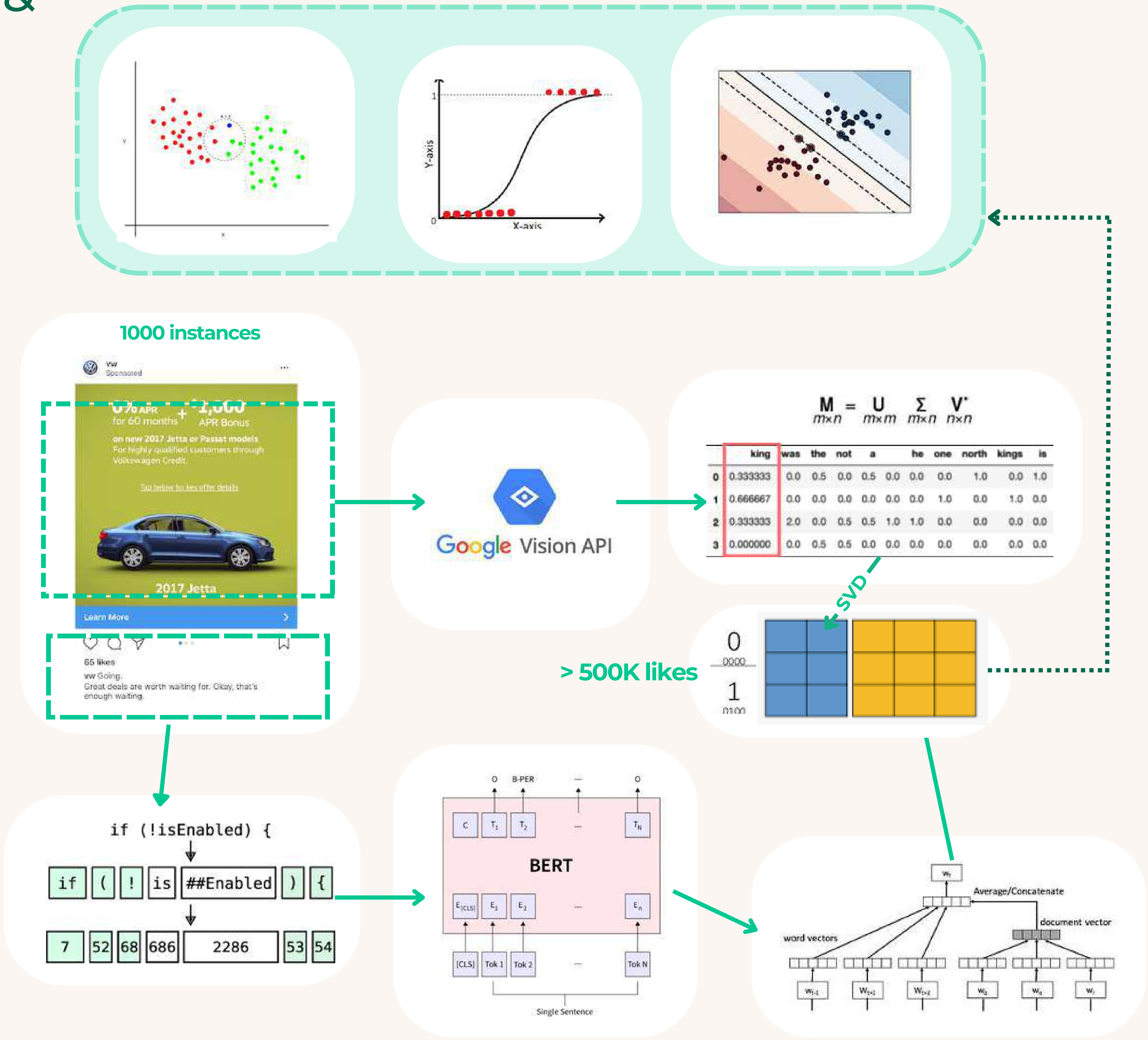
Challenge: Decoding effective combinations of text and imagery in ads to enhance user engagement and encourage conversions. Amidst the dynamic social media landscape, understanding these interactions is paramount for optimizing ad strategies.

Data Overview: Leveraging a diverse dataset sourced from Instagram, encompassing visual and textual elements across various influencers and themes, to analyze content engagement and refine ad strategies.

Machine Learning with Text Embedding & Image SVD on a **Small Dataset (1)**

Modeling | Part 1

Metrics	KNN	Logit Reg	SVM
In-sample Recall	0.31	1	1
In-sample Average Precision	0.32	1	0.99
In-sample Accuracy	0.93	1	0.99
Mean out-of-sample Recall	0.15	0.33	0.33
Mean out-of-sample Average Precision	0.17	0.2	0.19
Average out-of-sample Accuracy	0.9	0.89	0.88



McGill | DESAUTELS

Data Curation, Pre-Processing & Cleaning

Modeling | Parts 1 & 2

Data Acquisition

- **Data Source:** Instagram post images, images captions and meta data (likes) from different influencers scrapped with Python, complemented with Kaggle dataset.
- **Data Size: 5516 images from 1000 posts**
- ***Users: Influencers & businesses***

Data Pre-processing

- **Image Size:** All images resized to 128 x 128
- **Target: Whether a post has more than 500K likes (Binary Classification)**
- ***Split:***
  - ***4412 images for training dataset***
  - ***1104 images for test dataset***
  - ***883 images for validation dataset***

Data Wrangling & Preparation

- **Feature extraction:** Using Vertex AI = labels extraction, recognized in the image
- ***Scaling:***
  - ***Images scaled dividing by 255***
  - ***Features scaled using MinMax Scaler***

@mcgillu instagram

1080x1080

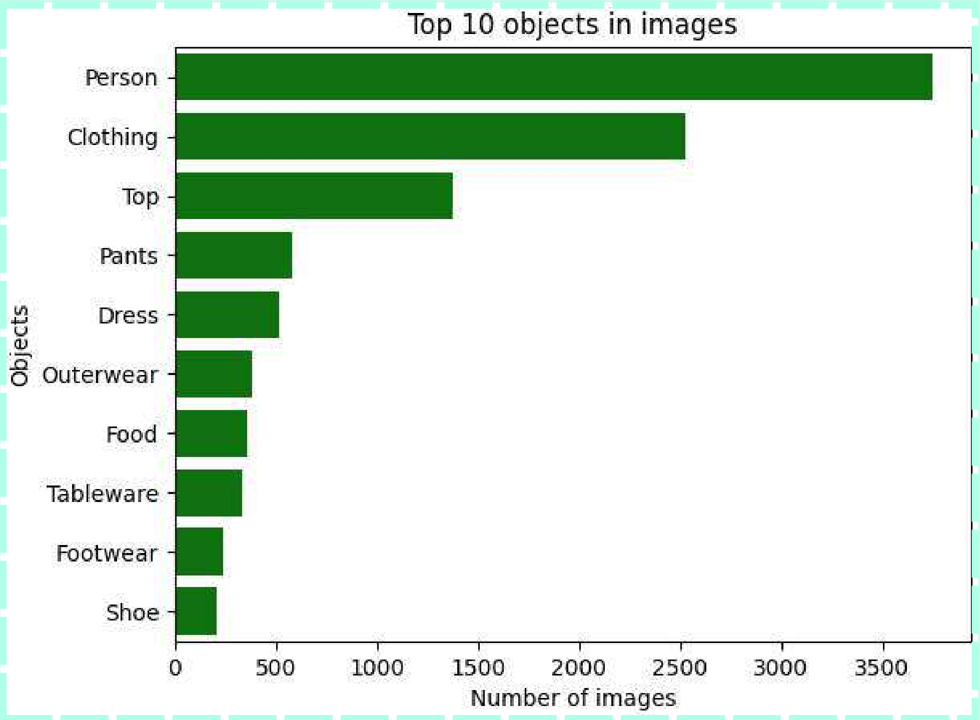
128x128

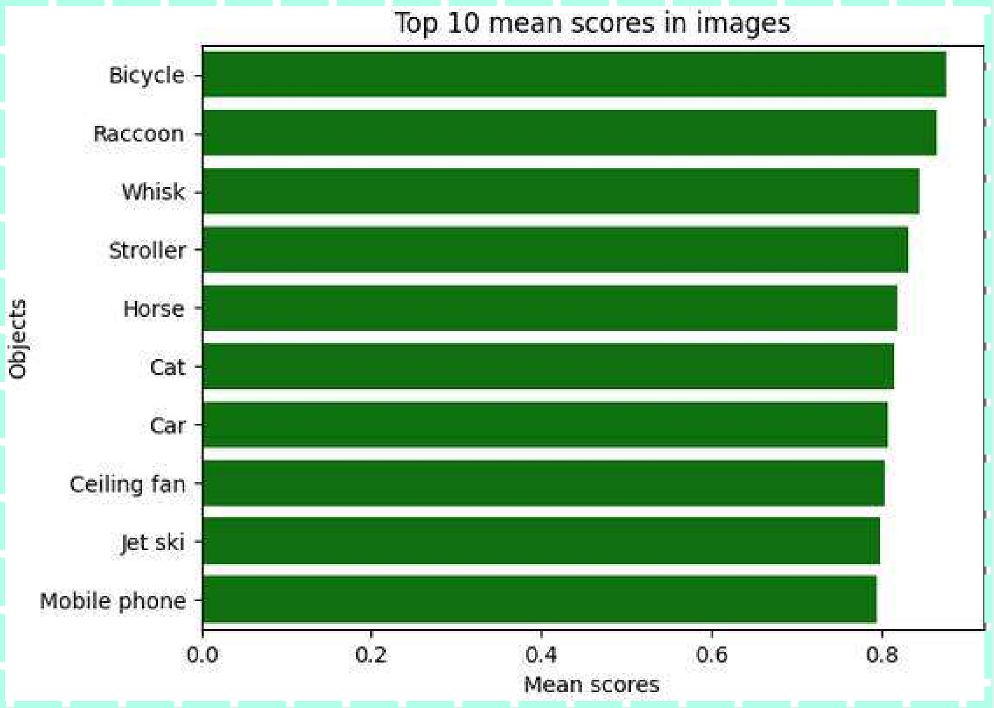
- **Table:** 0.99
- **Book:** 0.97
- **Person:** 0.99
- **Lamp:** 0.86

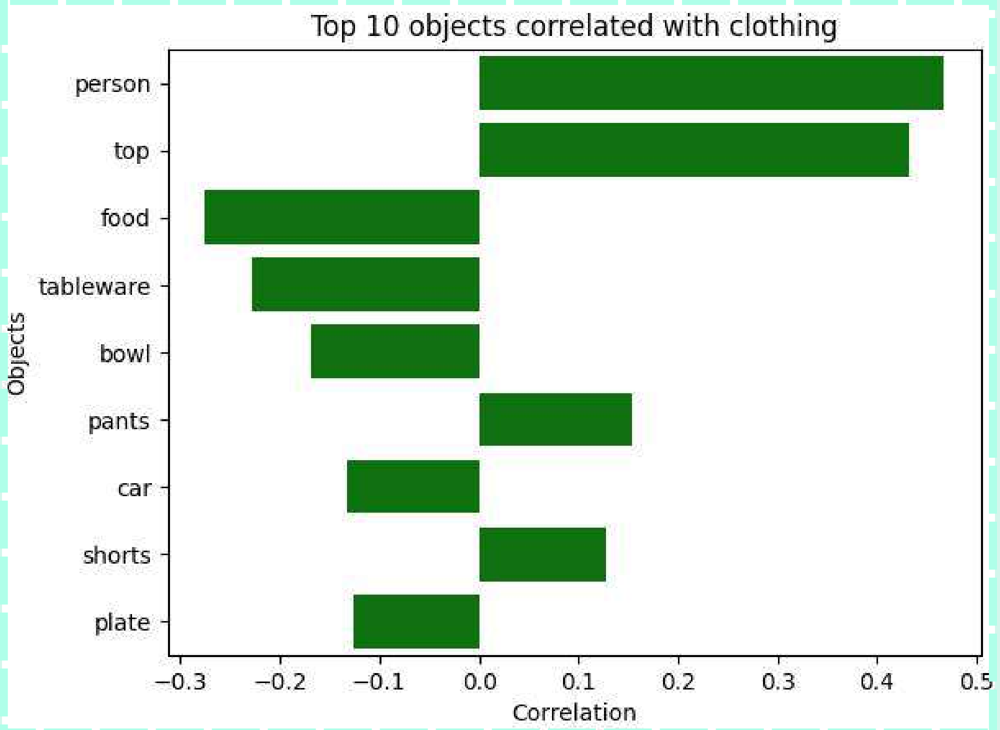
INSY 670 | Final Group Project

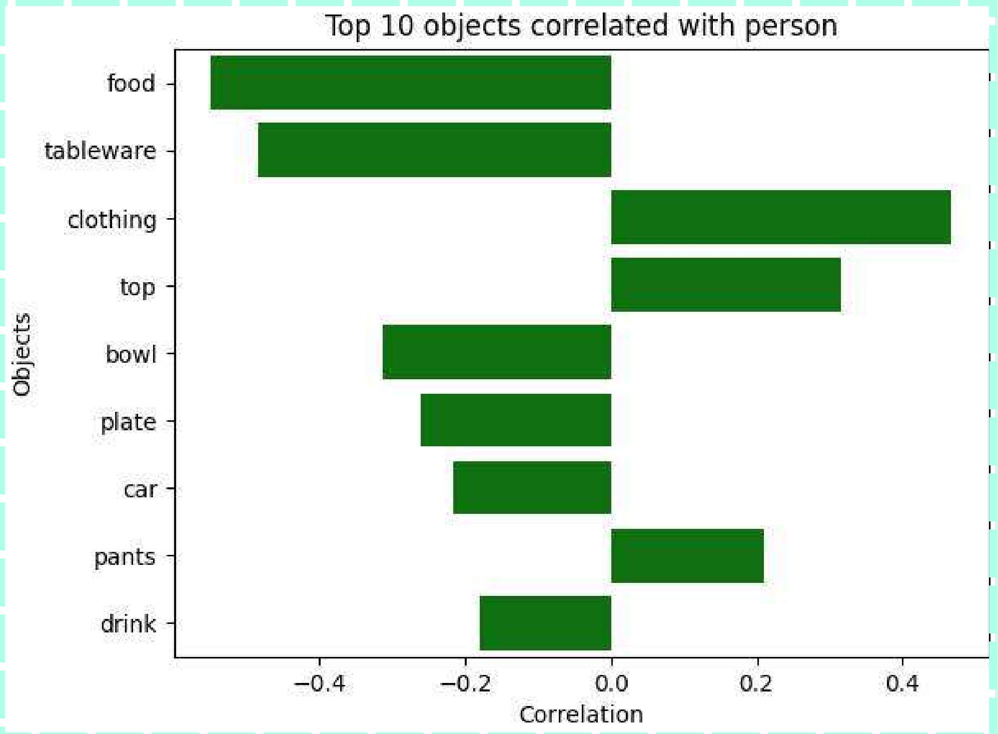
Data Curation Process & Preparation/Processing - Visualizing Insights

Modeling | Part 1



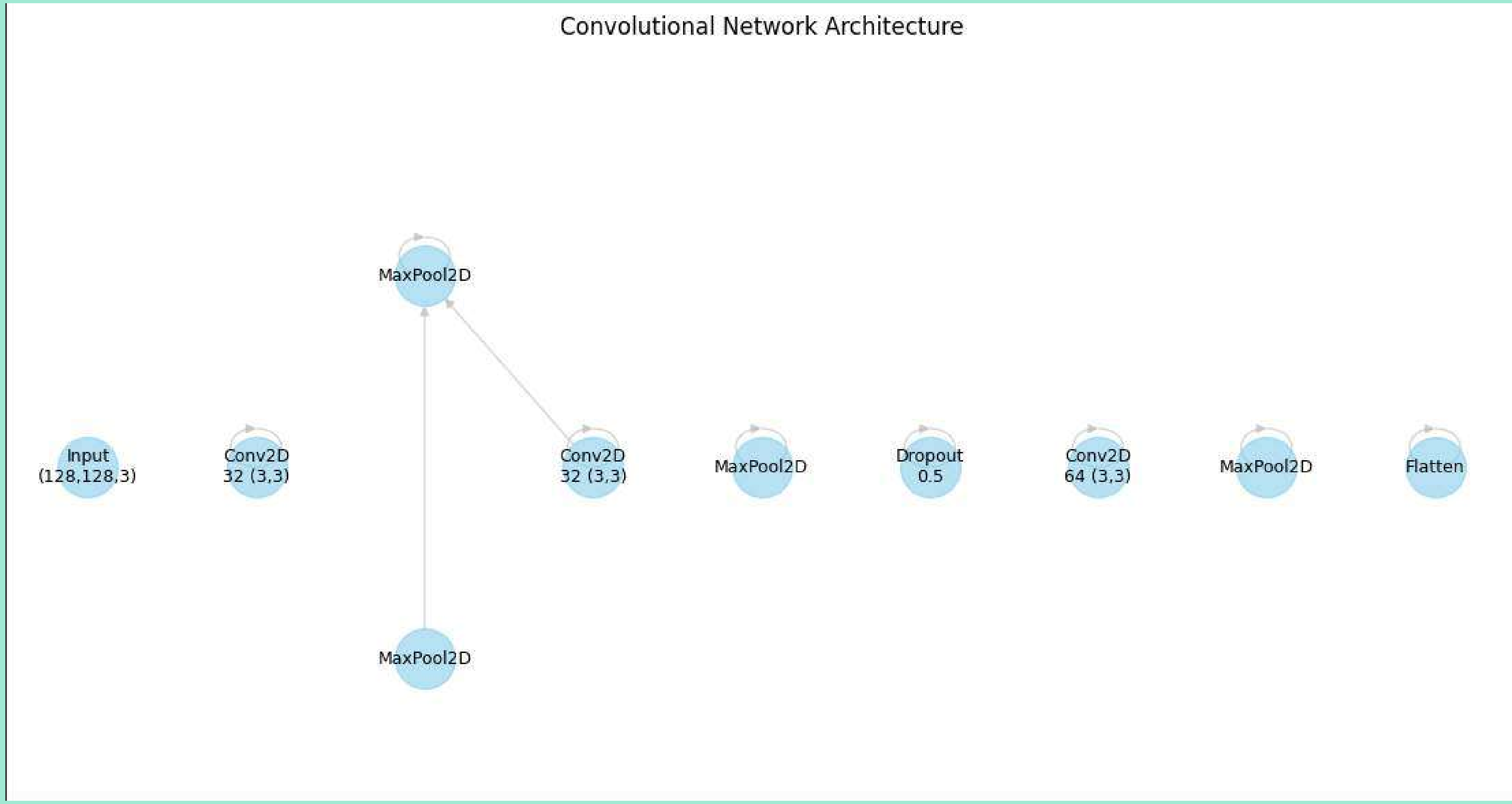




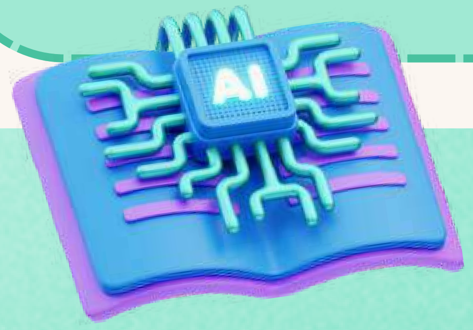


Combined Deep learning Model (CNNs & Perceptron) with a **Larger Dataset (2)**

Modeling | Part 2



Layer (type)	Output Shape	Param #	Connected to
input_layer (InputLayer)	(None, 128, 128, 3)	0	-
conv2d_1 (Conv2D)	(None, 126, 126, 32)	896	input_layer[0][0]
max_pooling2d_2 (MaxPooling2D)	(None, 63, 63, 32)	0	conv2d_1[0][0]
dropout (Dropout)	(None, 63, 63, 32)	0	max_pooling2d_2[…
conv2d_2 (Conv2D)	(None, 61, 61, 64)	18,496	dropout[0][0]
max_pooling2d_3 (MaxPooling2D)	(None, 30, 30, 64)	0	conv2d_2[0][0]
input_layer_1 (InputLayer)	(None, 255)	0	-
flatten (Flatten)	(None, 57600)	0	max_pooling2d_3[…
dense (Dense)	(None, 10)	2,560	input_layer_1[0]…
concatenate (Concatenate)	(None, 57610)	0	flatten[0][0], dense[0][0]
dense_1 (Dense)	(None, 256)	14,748,416	concatenate[0][0]
dense_2 (Dense)	(None, 1)	257	dense_1[0][0]



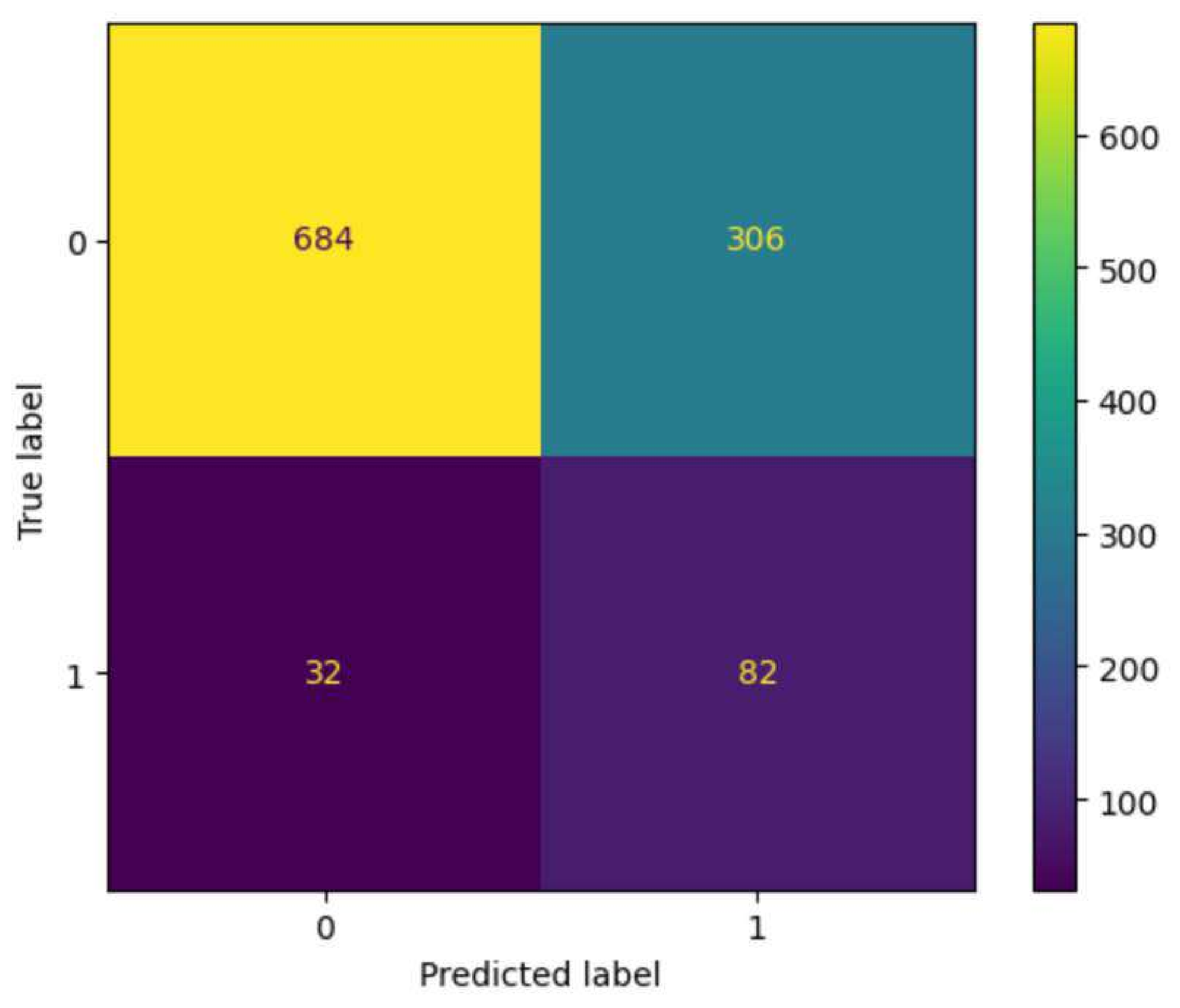
Analytical Approach & Methodology

/Demo/

McGill | DESAUTELS

Combined Deep learning Model (CNNs & Perceptron) with a **Larger Dataset (2)**

Modeling | Part 2 - Results & Insights



ROC-AUC Score (0.7429)

A score of 0.7429 indicates a good level of discriminative ability. The model has a relatively high likelihood of ranking a randomly chosen positive instance (more than 500K likes) higher than a randomly chosen negative instance (less than 500K likes).

Accuracy (0.6938)

An accuracy of 0.6938 means that about 69.38% of the predictions made by the model are correct.

Recall for Class 1 (0.7193)

A recall of 0.7193 for posts with more than 500K likes means that the model correctly identifies 71.93% of the high-like posts. This is quite a good result, indicating that the model has a strong ability to catch the high-performing posts but still misses some.

Recall for Class 0 (0.6909)

A recall of 0.6909 for posts with less than 500K likes means that the model correctly identifies 69.09% of the lower-performing posts.

Business Outcomes & Implications

1

Use of hybrid models for Social Media Ad Engagement Prediction

- Analysis of text and visual elements in Instagram ads
- Tailor brand's content management with LLMs and Deep Learning
- Identify key factors driving audience interaction.

2

Targeted Engagement: Precision Marketing

- Gain insights into platform users' interests and focal points
- Combine the main selling points with key attributes from prediction
- Stimulate substantial user interaction

3

Viral Potential Prediction: Enhancing ROI

- Access the potential to produce highly-liked post
- Help decrease the uncertainty in social media ad performance
- Enhance the expected ROI

4

Insights for Content Creators

- Help influencers understand hot topics and elements in viral content
- Cooperate with advertiser to generate higher engagement rates

McGill | DESAUTELS

Limitations & Future Steps

Limitations of Likes as Conversion Indicators

- A superficial engagement metric emphasizing entertainment over persuasiveness
- Indicating audience interest without directly measuring purchasing power
- Important to consider the objectives of the campaign
  - *For campaigns aiming to expand brand influence, focusing on likes is reasonable*

Variability in Engagement Across Different Sectors

- Varying base sizes of target audiences and distinct user behavior patterns
- Setting varying thresholds for likes to provide more detailed predictions

Lack of Negative Feedback Collection

- The importance of multifaceted feedback from campaigns in the long term
- Highlighting the aspects to focus on with the help of positive feedbacks
- Emphasizing what to avoid incorporating the collection of negative feedback

More Attributes than Image Features

- Integrating additional types of attributes in the future
  - *Attributes related to influencers, networks, etc.*

Any Questions?

Thank you!