CAPSTONE I

Using Natural Language Processing and Computer Vision for Disaster Search and Rescue

Project by
Izza Godinez (100556078)
Michael Molnar (100806823)
Hibba Imtiaz (100794061)
Utsav Vanodiya (100804273)

Final Report

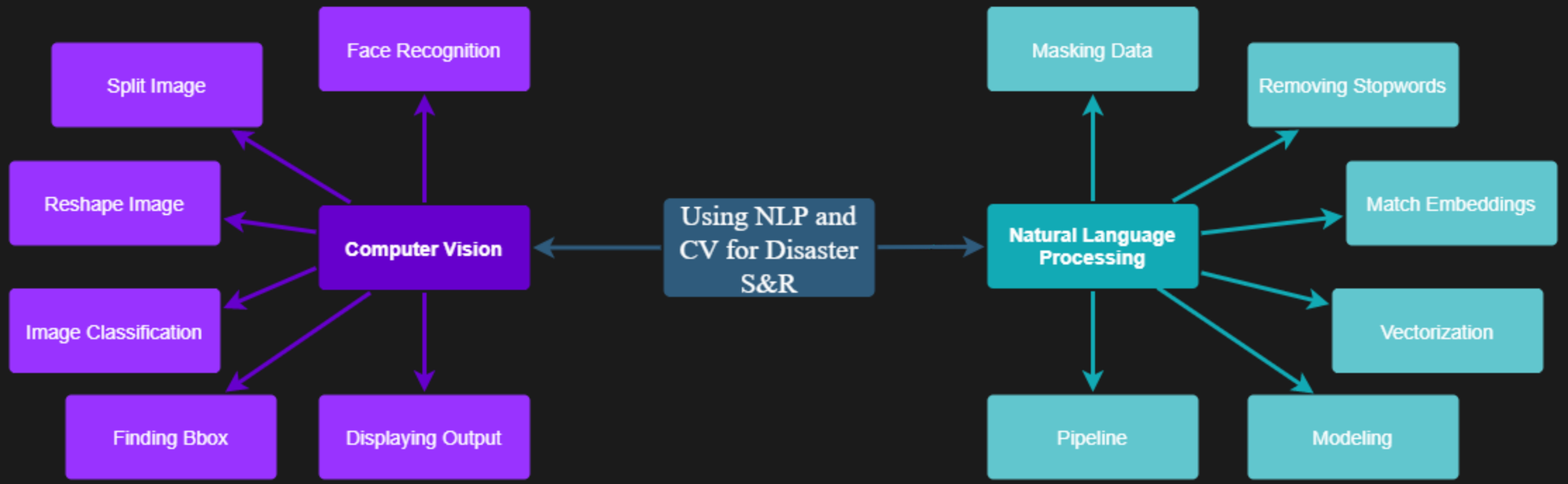
Final
Report

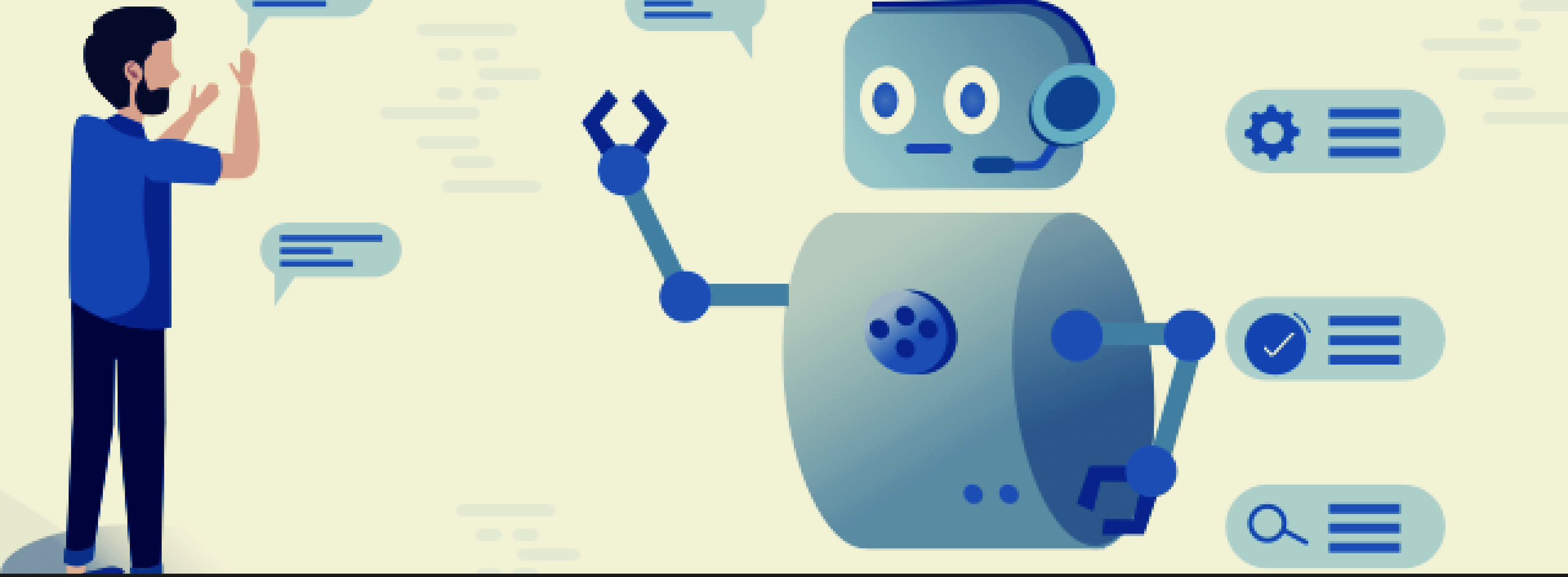
Contents

USING NATURAL LANGUAGE PROCESSING AND COMPUTER VISION FOR DISASTER RESCUE

INTRODUCTION

APPROACH


Face Recognition
Split Image
Reshape Image
Computer Vision
Image Classification
Finding Bbox
Displaying Output
Using NLP and CV for Disaster S&R
Masking Data
Removing Stopwords
Match Embeddings
Natural Language Processing
Vectorization
Pipeline
Modeling



Part 1
NLP Extracting Clothing Descriptors from Text

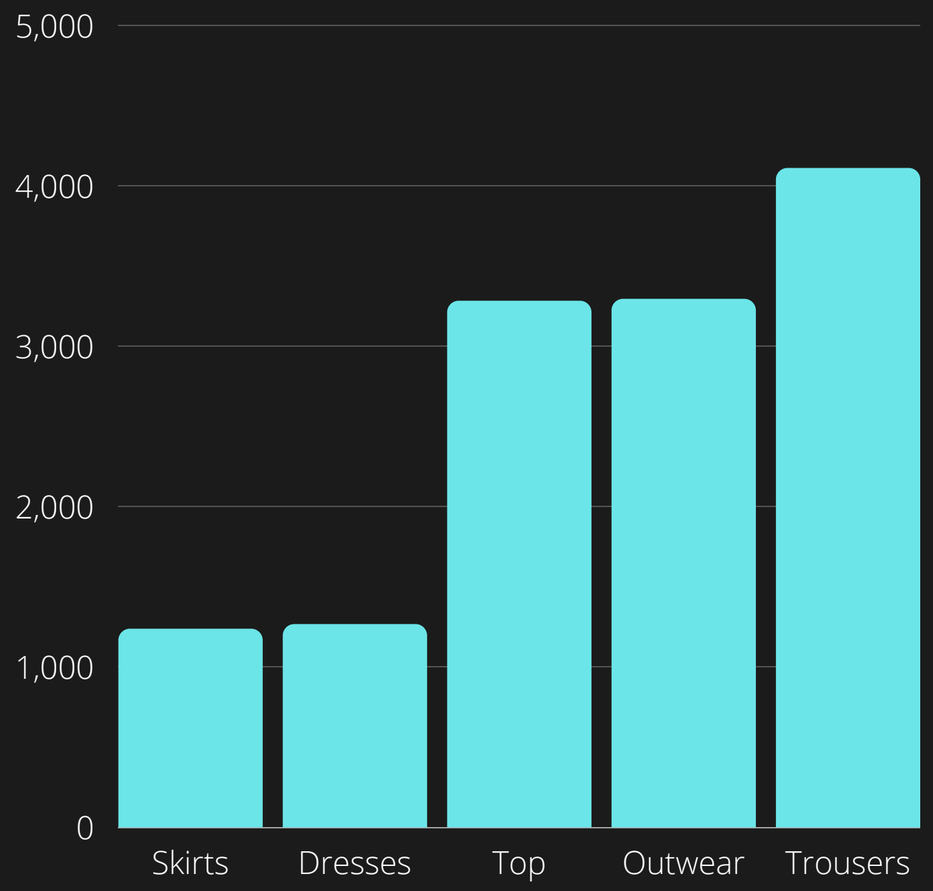
Final Report

Exploratory Data Analysis

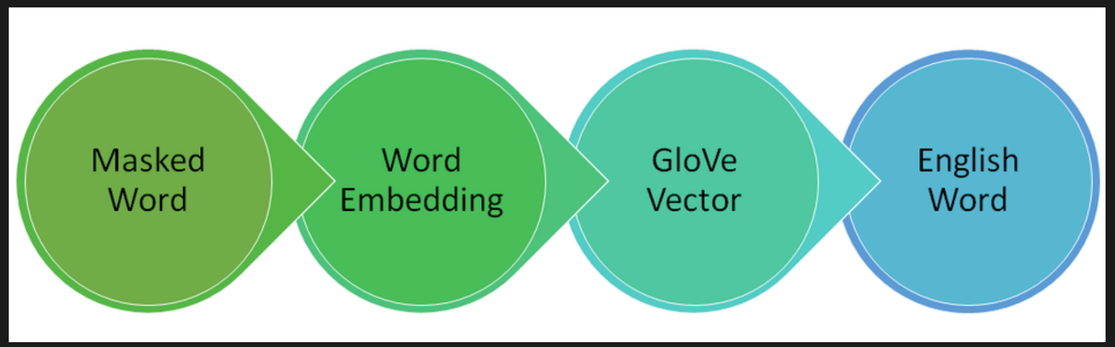
Number of rows = 7380
Number of columns = 7

Sample of data:

	id	word_representation	outwear	top	trousers	women dresses	women skirts
0	0	w7718 w173355 w138132 w232277 w90685 w314686 w...	1	0	1	0	0
1	1	w195317 w127737 w171593 w22890 w342007 w217871...	1	0	1	0	0
2	2	w247655 w270233 w261113 w337250 w366000 w37873...	0	1	1	0	0
3	3	w279289 w395855 w61795 w286461 w308610 w27013 ...	1	0	1	0	0
4	4	w254516 w135431 w115724 w331534 w256214 w71240...	1	0	1	0	0



Data Pre-processing

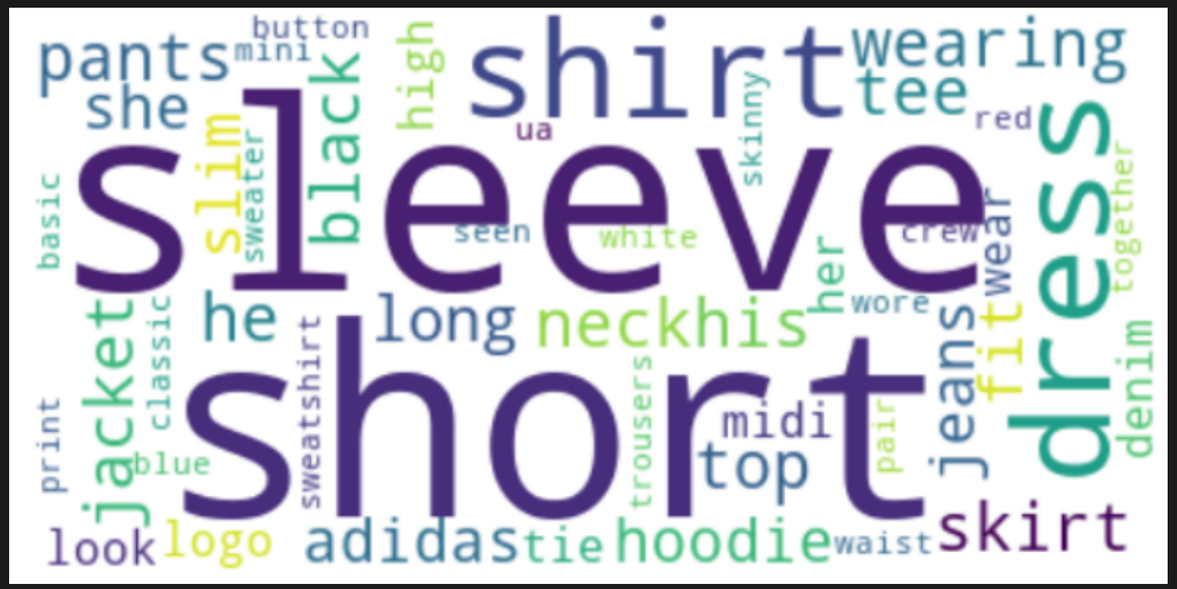


“w194870”

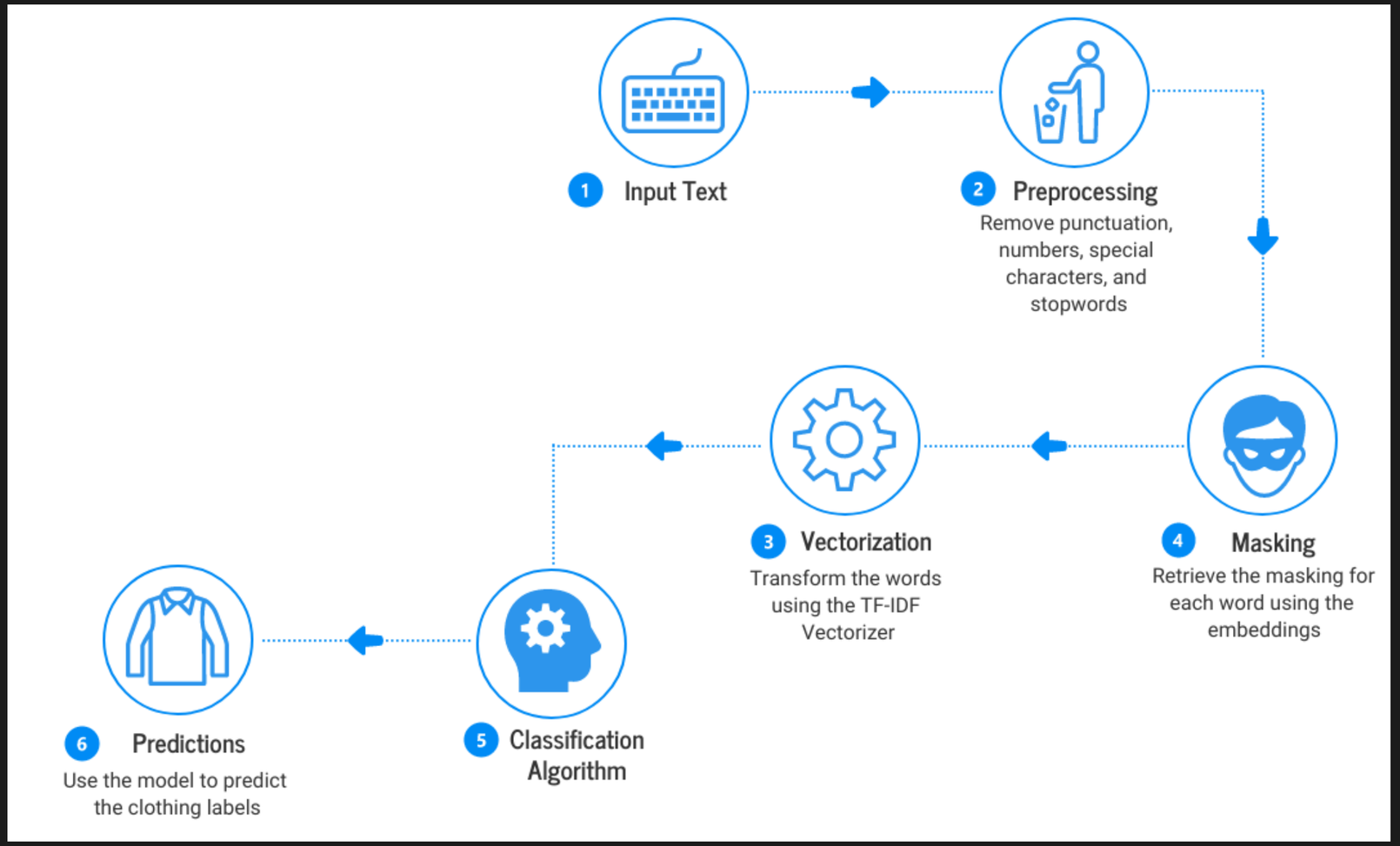

```
['-0.038194' '-0.24487' '0.72812' '-0.39961' '0.083172' '0.043953'
 '-0.39141' '0.3344' '-0.57545' '0.087459' '0.28787' '-0.06731' '0.30906'
 '-0.26384' '-0.13231' '-0.20757' '0.33395' '-0.33848' '-0.31743'
 '-0.48336' '0.1464' '-0.37304' '0.34577' '0.052041' '0.44946' '-0.46971'
 '0.02628' '-0.54155' '-0.15518' '-0.14107' '-0.039722' '0.28277'
 '0.14393' '0.23464' '-0.31021' '0.086173' '0.20397' '0.52624' '0.17164'
 '-0.082378' '-0.71787' '-0.41531' '0.20335' '-0.12763' '0.41367'
 '0.55187' '0.57908' '-0.33477' '-0.36559' '-0.54857' '-0.062892'
 '0.26584' '0.30205' '0.99775' '-0.80481' '-3.0243' '0.01254' '-0.36942'
 '2.2167' '0.72201' '-0.24978' '0.92136' '0.034514' '0.46745' '1.1079'
 '-0.19358' '-0.074575' '0.23353' '-0.052062' '-0.22044' '0.057162'
 '-0.15806' '-0.30798' '-0.41625' '0.37972' '0.15006' '-0.53212' '-0.2055'
 '-1.2526' '0.071624' '0.70565' '0.49744' '-0.42063' '0.26148' '-1.538'
 '-0.30223' '-0.073438' '-0.28312' '0.37104' '-0.25217' '0.016215'
 '-0.017099' '-0.38984' '0.87424' '-0.72569' '-0.51058' '-0.52028'
 '-0.1459' '0.8278' '0.27062']
```


“the”

Stopword Removal



Architecture Diagram



Vectorization TF-IDF

df_tfidfvect

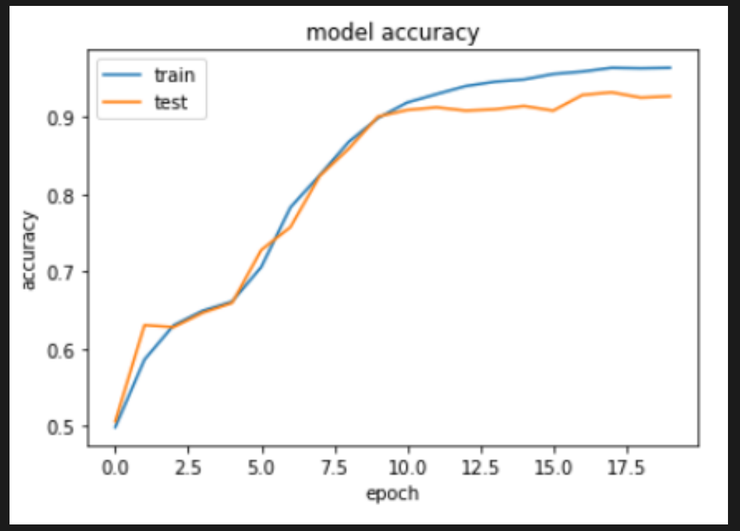
	w100060	w1001	w100157	w100187	w100269	w100299	w100527	w10065	w100799	w100966	...	w99014	w99144	w99304	w99321	w99479	w99485
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
5161	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
5162	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
5163	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
5164	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
5165	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0

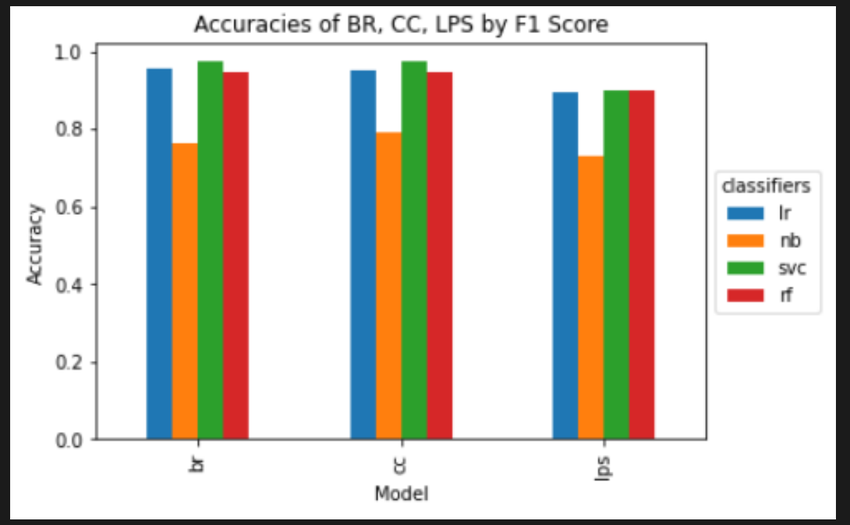
5166 rows × 3716 columns

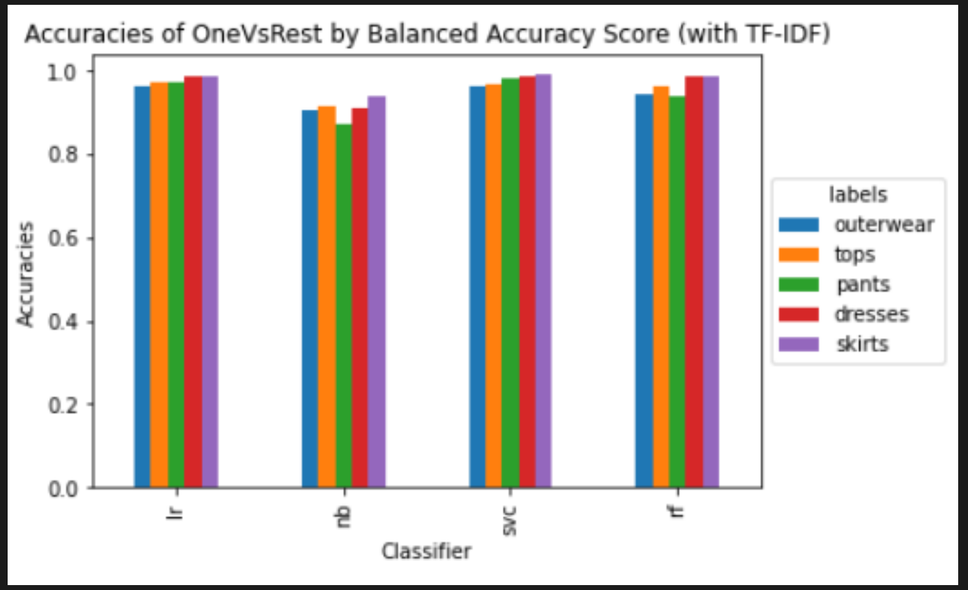
5166 training rows, 3716 unique words

Chosen Algorithm

OneVsRest with a Logistic Regression Classifier

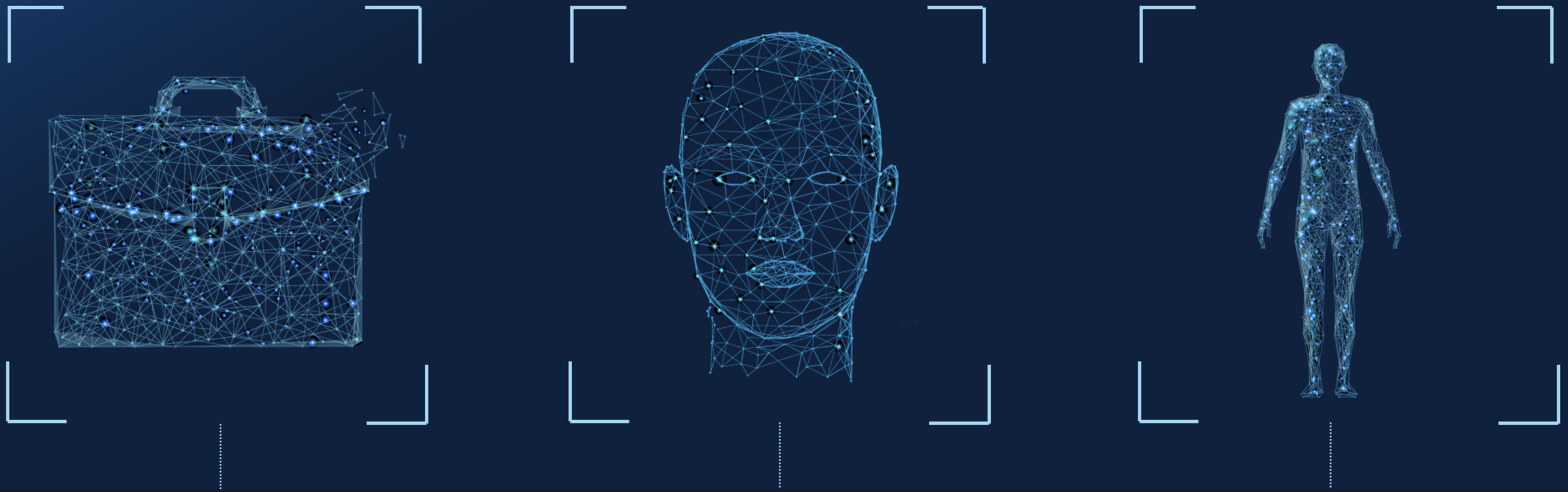






Live Demo

```
predict_labels()
```



Part 2: Computer Vision Detecting Clothing Objects

Final Report

Exploratory Data Analysis

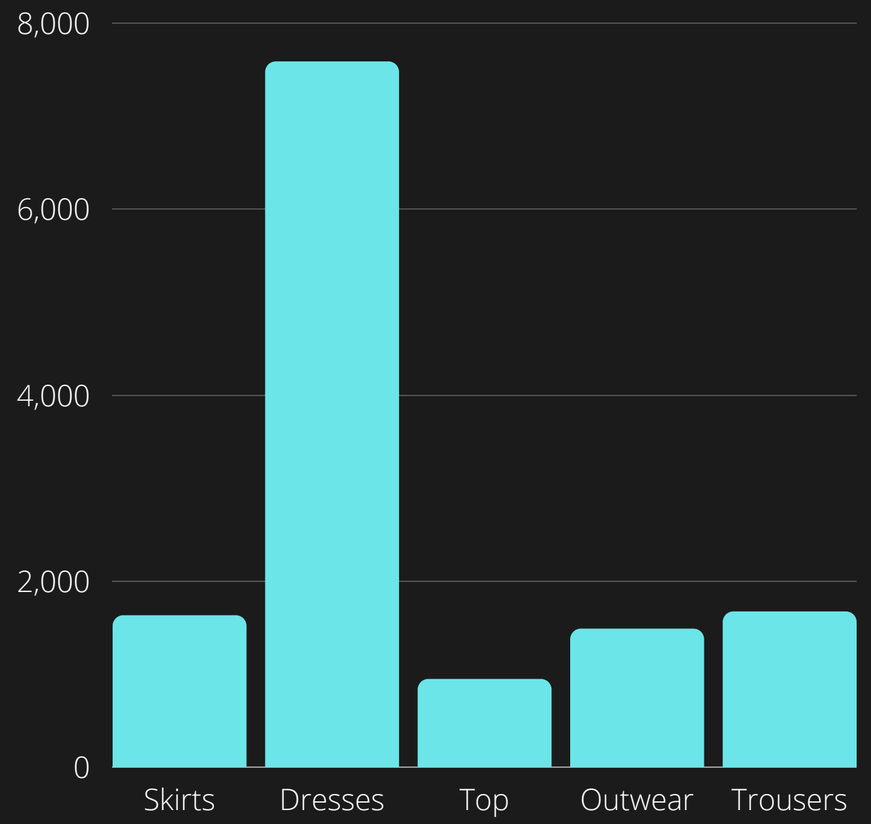
JSON example:
"images":[{ "file_name": "1.jpg" "id": "1" }]"

annotations":[{ "id": 1, "image_id": 10, "category_id": 4, "bbox": [704, 620, 1401, 1645] # left, top, width, height (xywh) }]"

categories": [{"id": 1, "name": "tops"}, {"id": 2, "name": "trousers"}, {"id": 3, "name": "outerwear"}, {"id": 4, "name": "dresses"}, {"id": 5, "name": "skirts"}]}

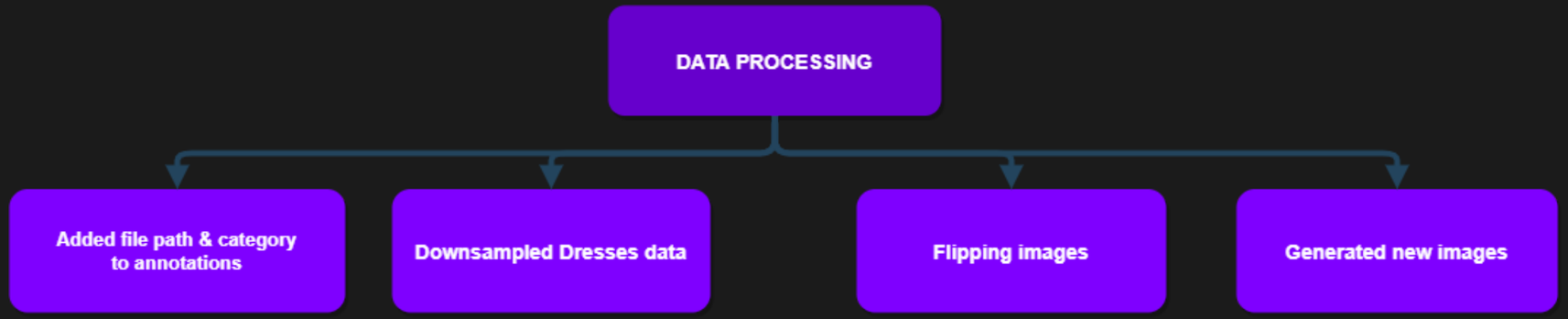
Number of training images: 8,225
Number of test images: 1,474

Tops: 945
Trousers: 1671
Outwear: 1486
Dresses: 7585
Skirt: 1630

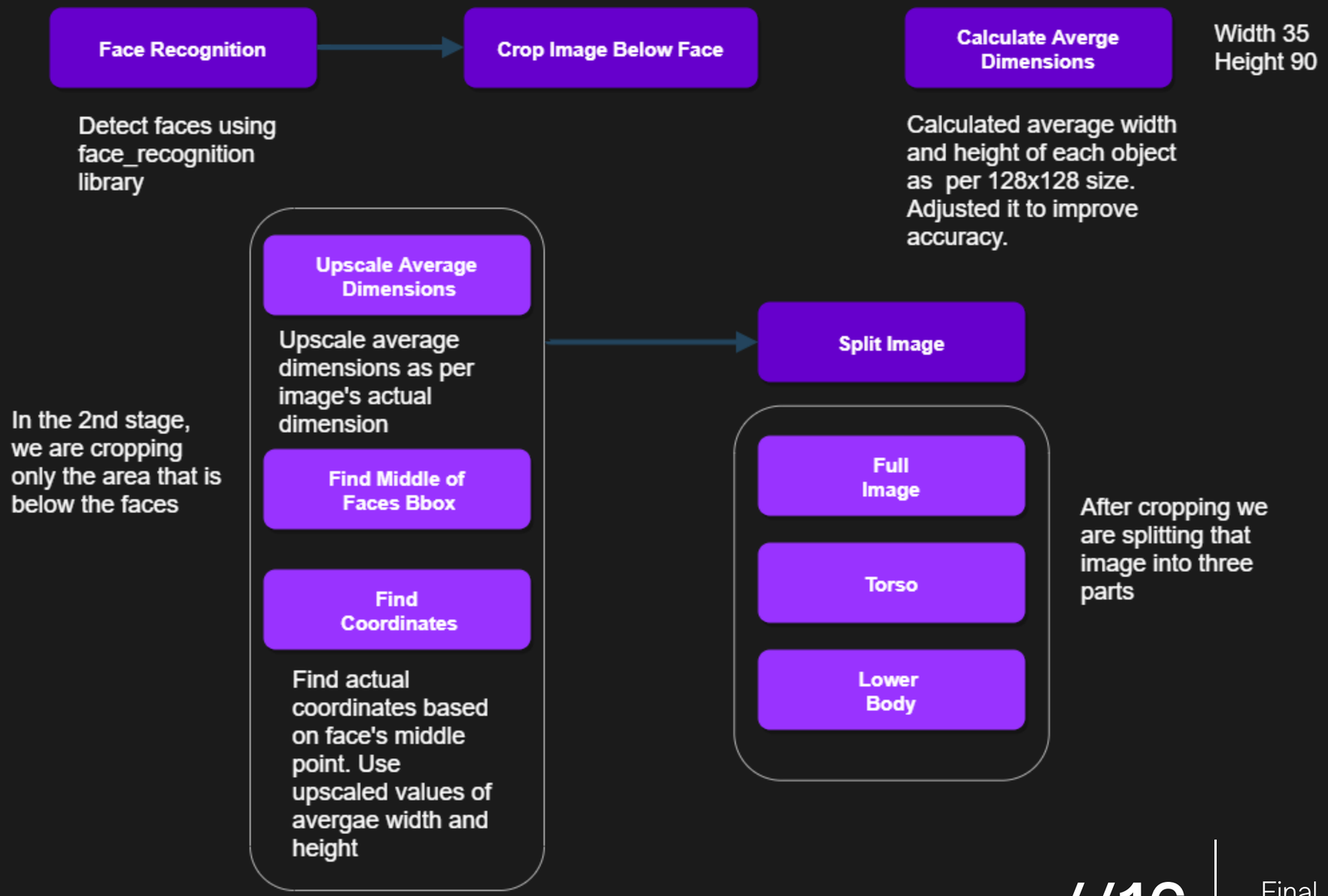


Data Pre-processing

- Added file path and category to annotations.
- Downsampled Dresses data.
- Generated new images for other categories by flipping images.
- Generated new images for the noise category. So, when face detection fails, we need to scan entire image and that time this category will become helpful.



Object Detection



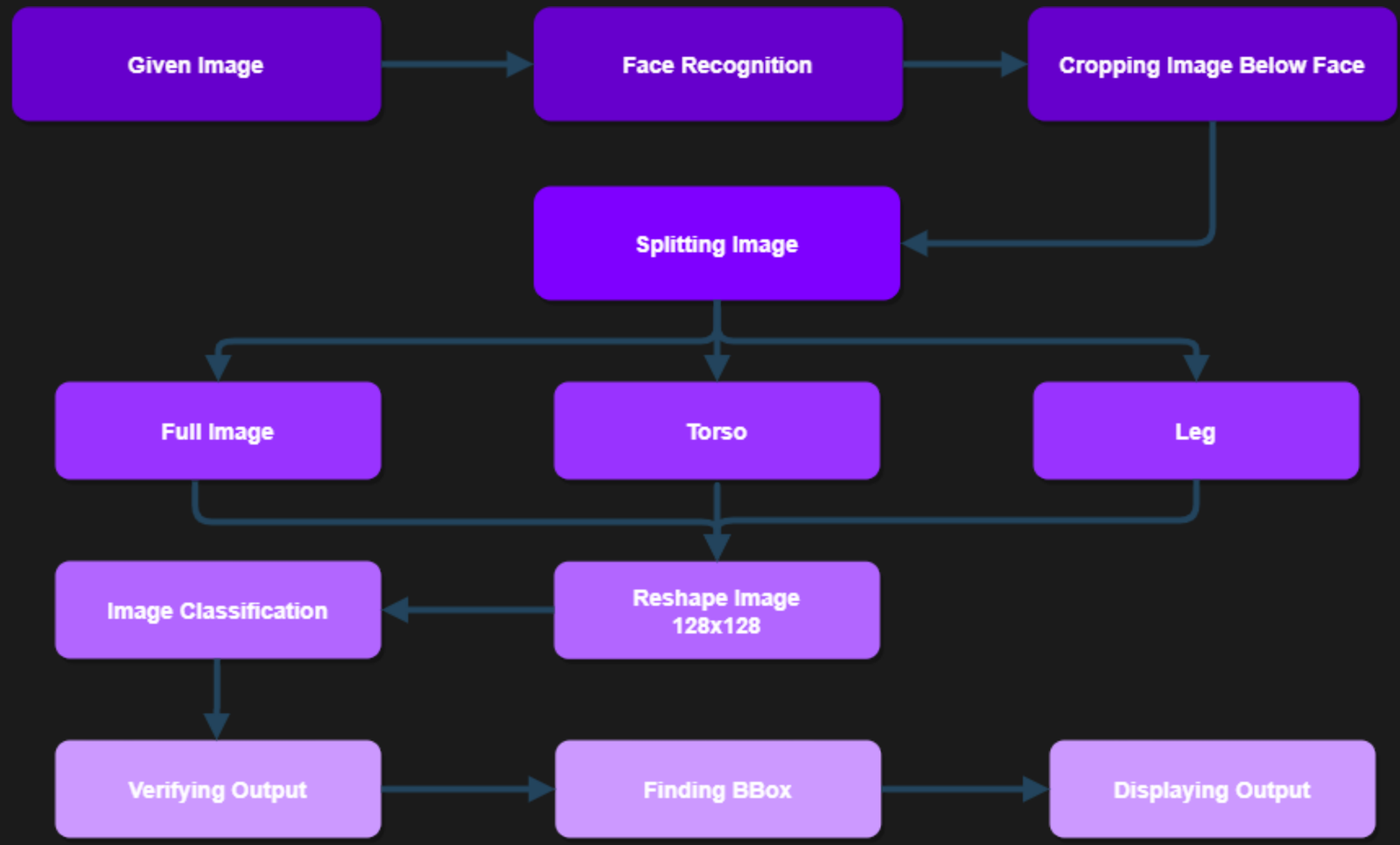
Example of Cropping & Splitting Image

Actual Image

Cropped Image

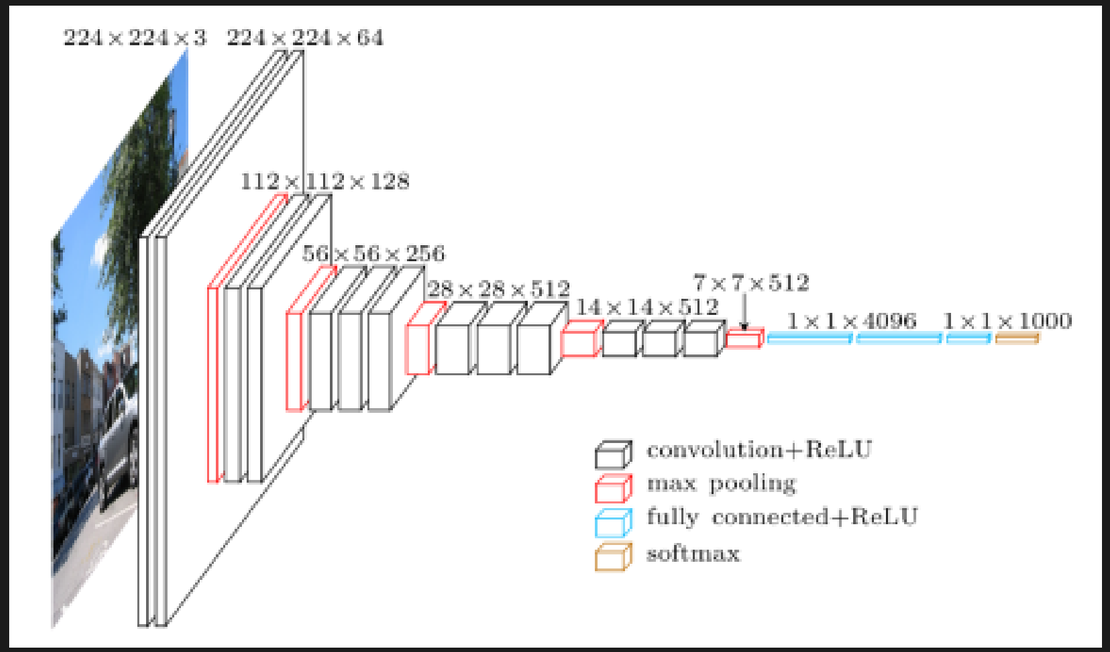
Split Cropped Image into 3 parts

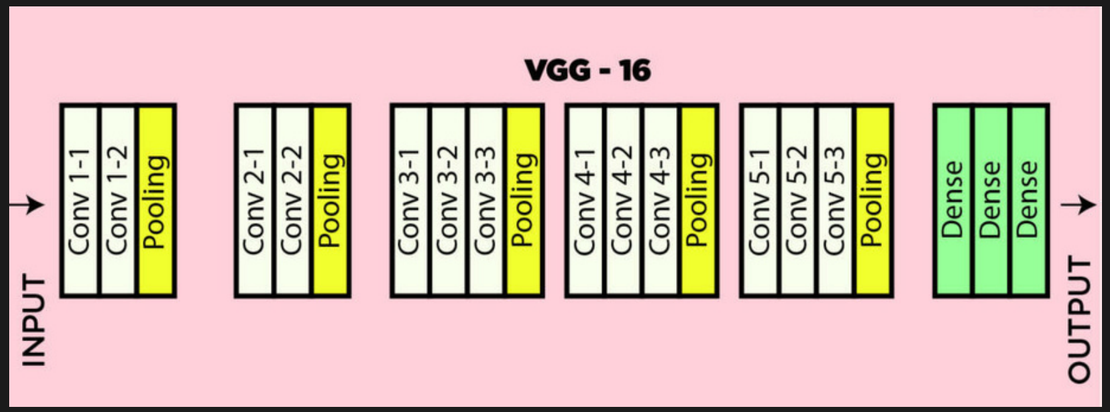
Architecture Diagram

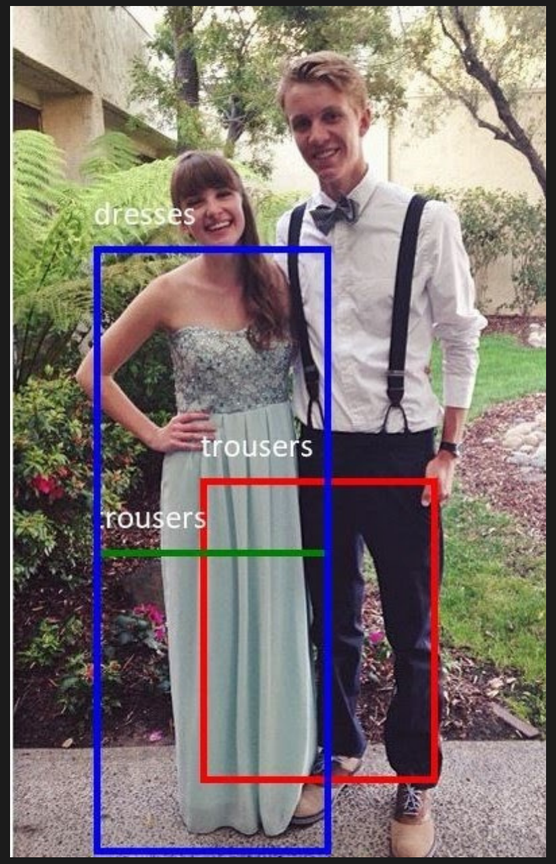


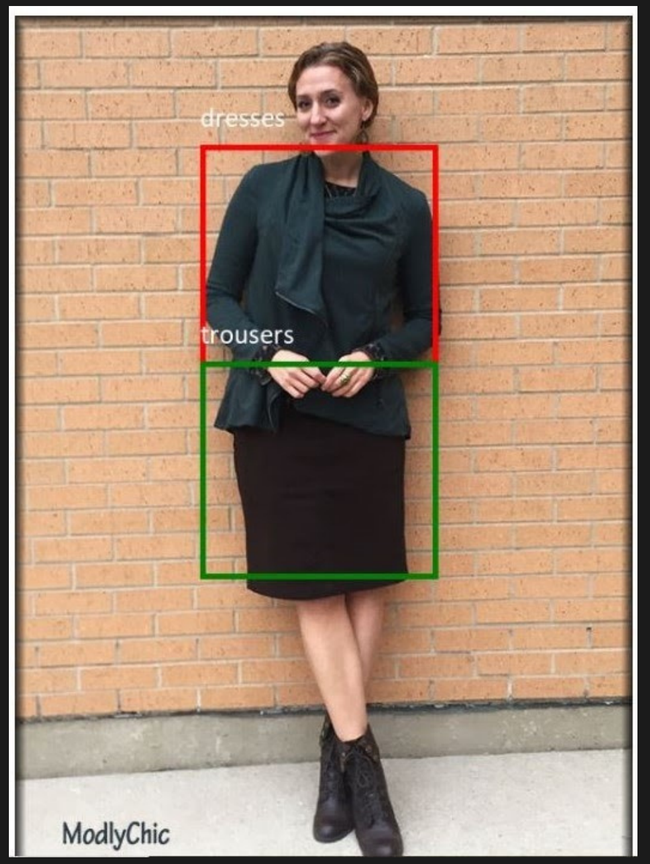
Chosen Algorithm

VGG16 to train an Image Classifier









Output Examples

Live Demo

Conclusion

Final
Report

Thank You