Deepfake Detection

By

MEDHINENI AMULYA – 99220041832

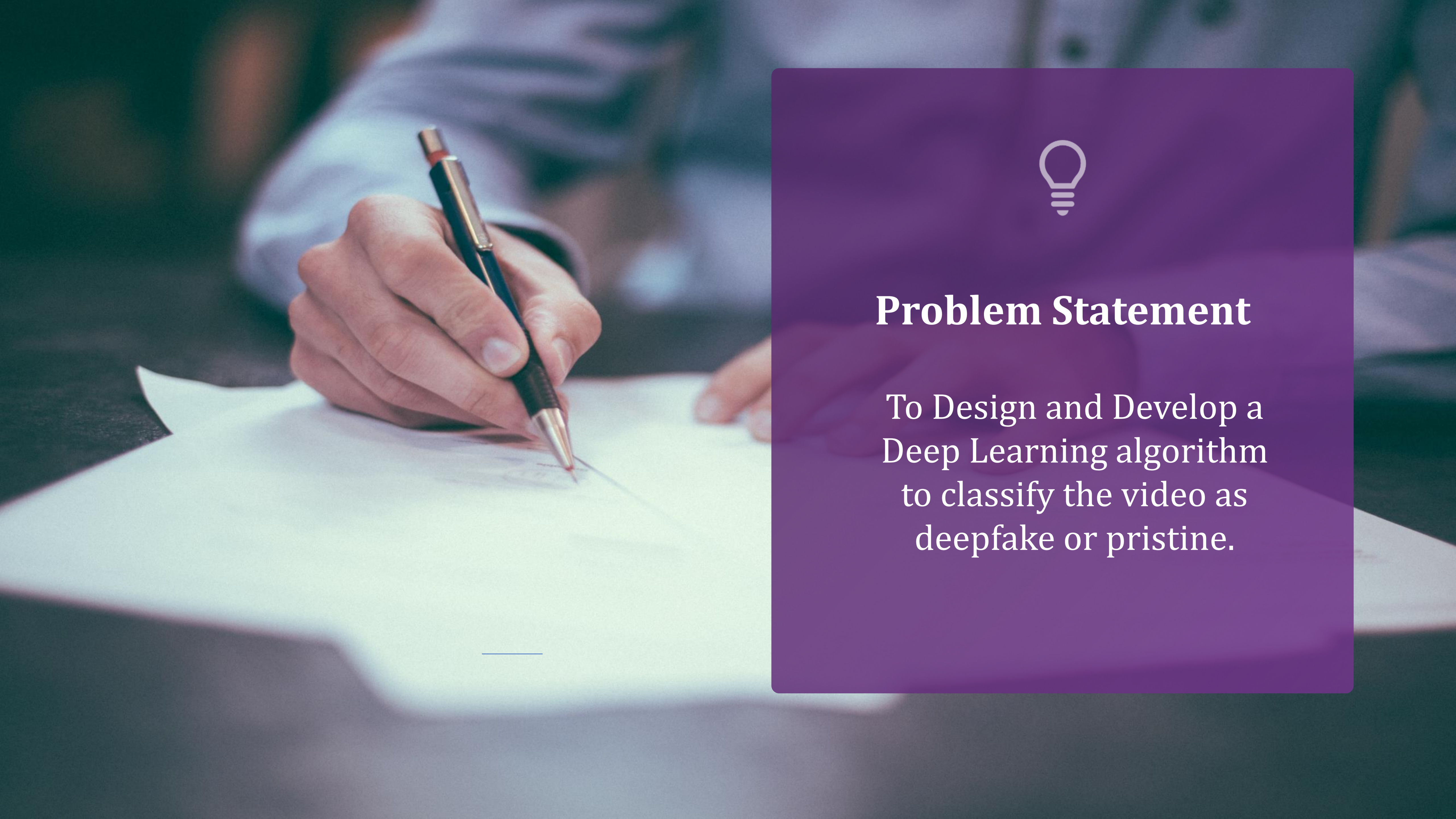
M SOWMYA – 99220041838

B.MANIKANTA – 99220041819

A.PRASHANTH KUMAR - 99220042062

Guided by:

MRs.K.BAVANI KANNAN



Problem Statement

To Design and Develop a Deep Learning algorithm to classify the video as deepfake or pristine.

01 Introduction

02 System Architecture

03 Dataset Exploration

04 Pre-processing

05 Model Architecture

06 Training Workflow

07 Prediction Workflow

08 Tools and Technologies

09 Results

10 Demo

❑ Introduction

- Deep fake is a technique for human image synthesis based on artificial intelligence.
- Deep fakes are created by combing and superimposing existing images and videos onto source images or videos using a deep learning technique known as generative adversarial network.

❑Can we detect Deep fakes with naked eyes?

GJ

Real or Fake?

عربية

غرينتش 14:01
ل عبور الحالات الإنسانية
إسرائيل تفرج عن الناشط صدقي المقت
From a previous episode

Real

Fake

Fake

Fake

[Rössler et al. 2019]

Source

Fake

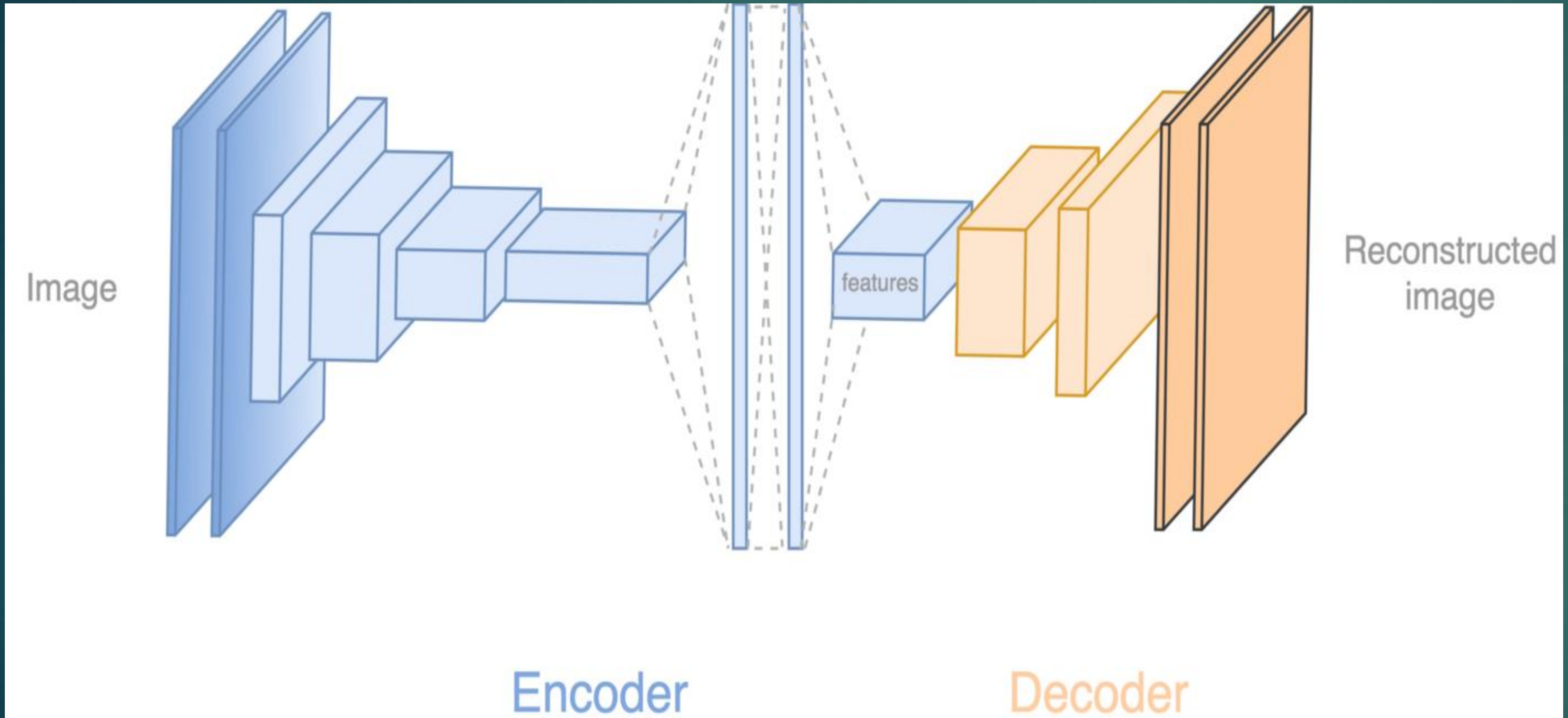
غرينتش 14:01 الحالات الإنسانية إسرائيل تفرج عن الناشط صدقي المقت بعد 27

Real

Why Deep Fake Detection ?

- Fake News
- Malicious hoaxes
- Financial fraud
- Celebrity unusual video
- Revenge porn
- Politician videos

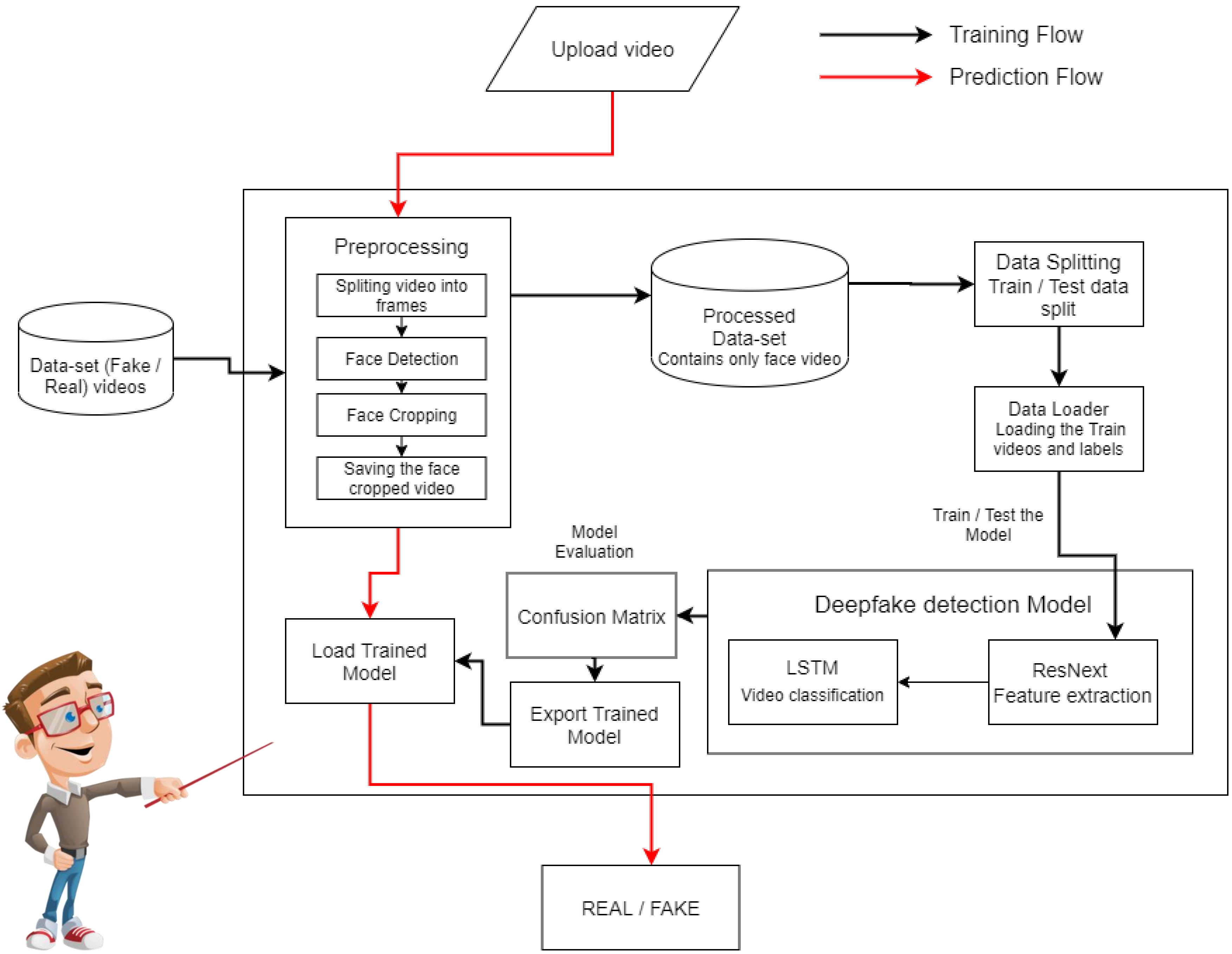
❑ How Deep Fakes Are Created ?



Tools for deep fake creation.

- Faceswap
- Faceit
- DeepFaceLab
- DeepfakeCapsuleGAN
- Large resolution facemasked

System Architecture



Upload video

Training Flow

Prediction Flow

Data-set (Fake / Real) videos

Preprocessing

Spliting video into frames

Face Detection

Face Cropping

Saving the face cropped video

Processed Data-set Contains only face video

Data Splitting Train / Test data split

Data Loader Loading the Train videos and labels

Train / Test the Model

Model Evaluation

Deepfake detection Model

ResNext Feature extraction

LSTM Video classification

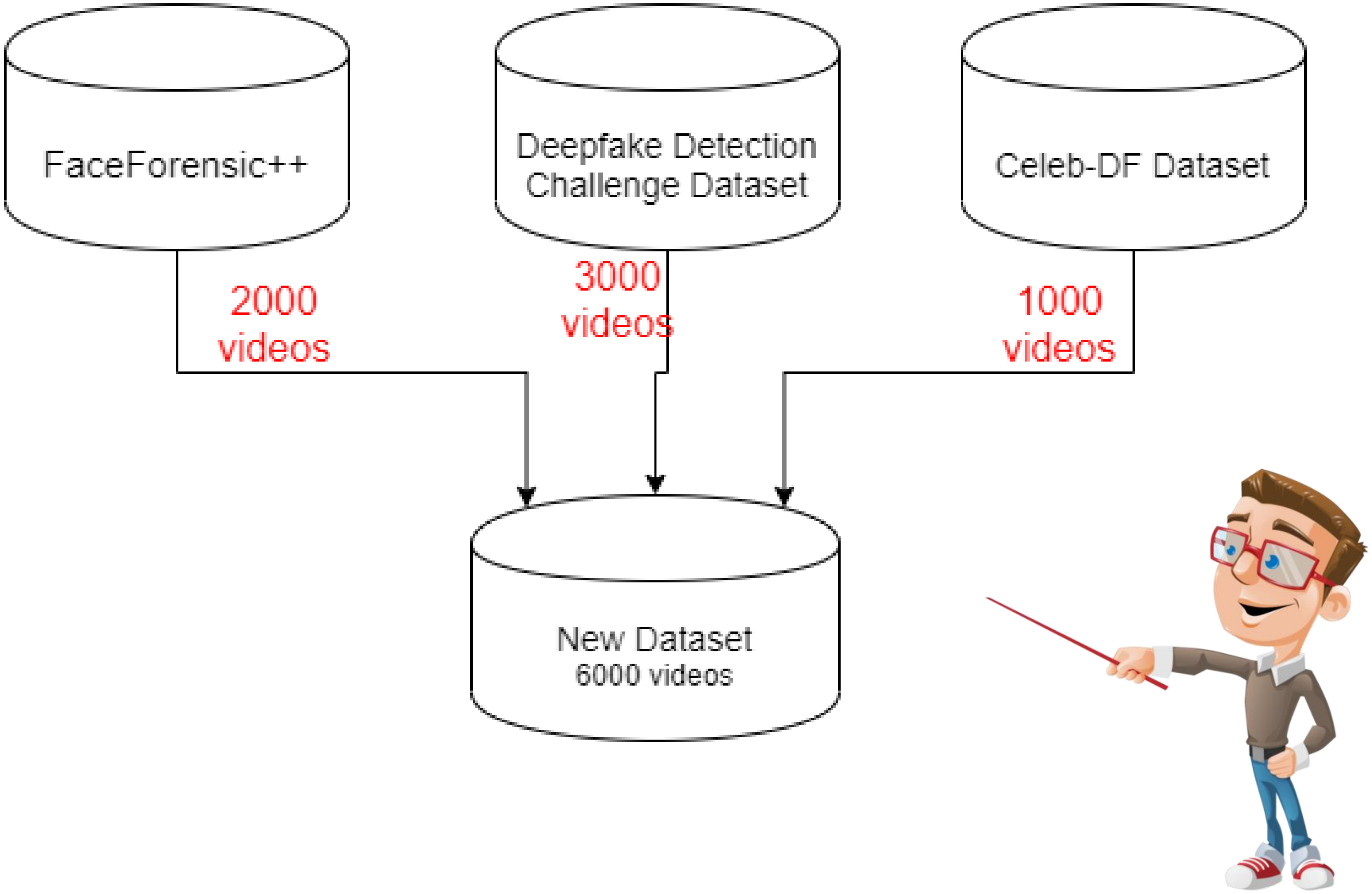
Confusion Matrix

Export Trained Model

Load Trained Model

REAL / FAKE

Data-set Exploration



FaceForensic++

Deepfake Detection Challenge Dataset

Celeb-DF Dataset

2000 videos

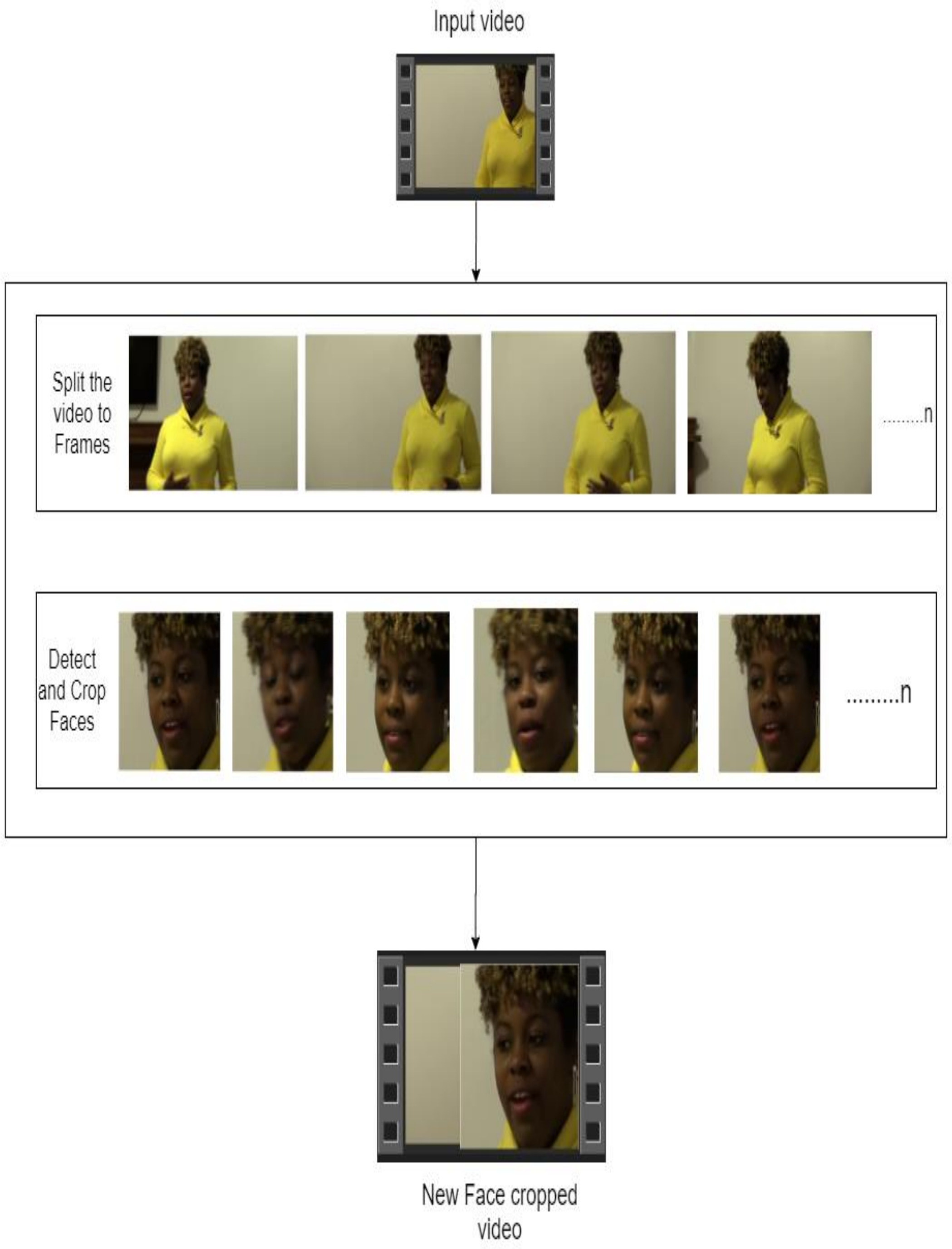
3000 videos

1000 videos

New Dataset
6000 videos

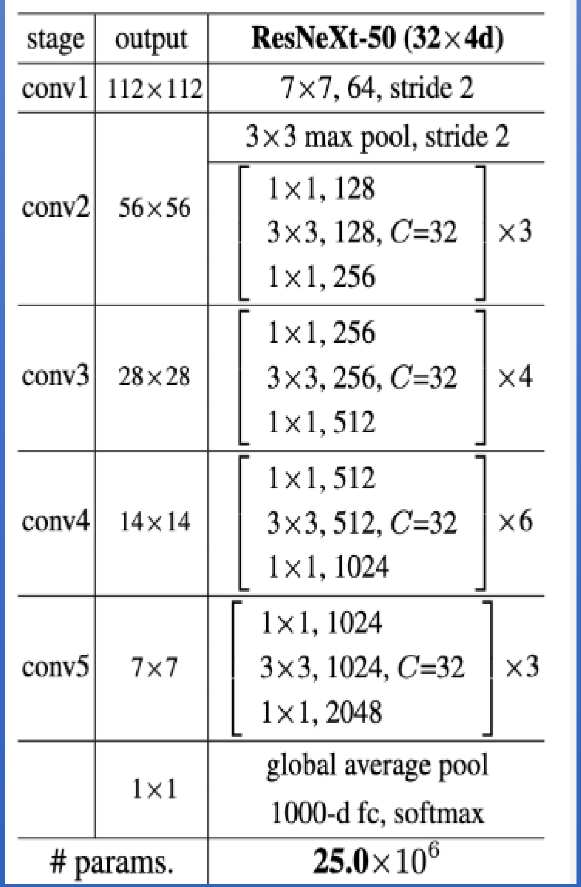
Pre-processing

1. SPLIT VIDEO INTO FRAMES
2. FACE DETECTION
3. CROPING FACE
4. CREATING NEW FACE CROPPED VIDEO
5. SAVING THE FACE CROPPED VIDEO



Model Architecture

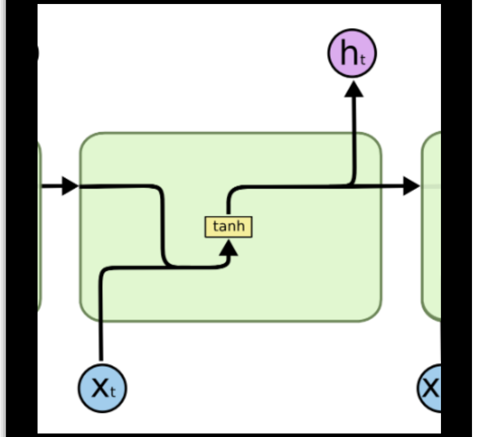
ResNext-50

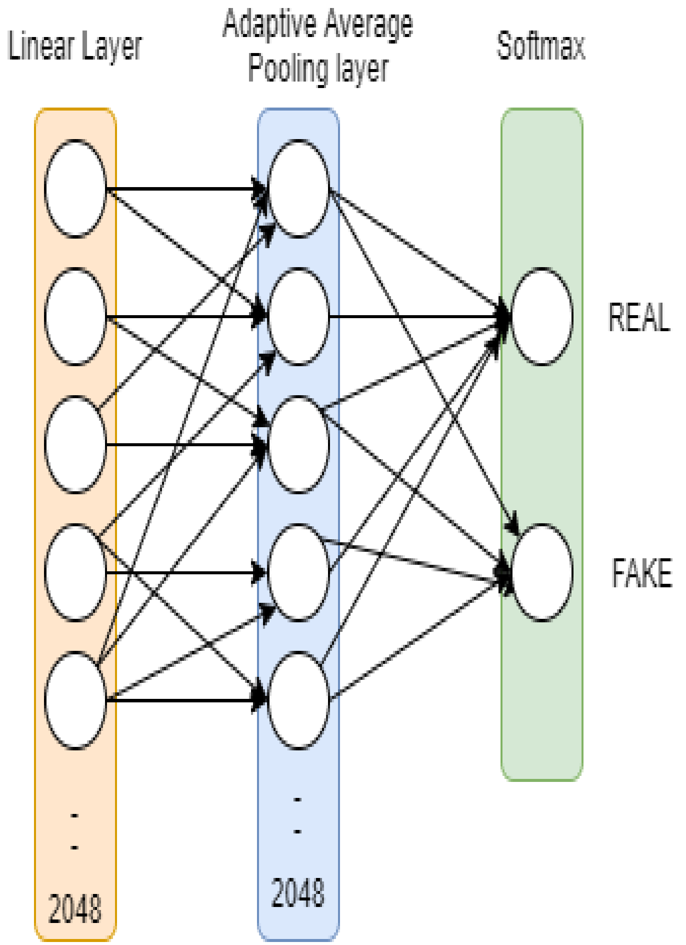


stage	output	**ResNeXt-50 (32×4d)**
conv1	112×112	7×7, 64, stride 2
conv2	56×56	3×3 max pool, stride 2 $\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128, C{=}32 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3	28×28	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256, C{=}32 \\ 1\times1, 512 \end{bmatrix}\times4$
conv4	14×14	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512, C{=}32 \\ 1\times1, 1024 \end{bmatrix}\times6$
conv5	7×7	$\begin{bmatrix} 1\times1, 1024 \\ 3\times3, 1024, C{=}32 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	global average pool 1000-d fc, softmax
# params.		**25.0**×10^6

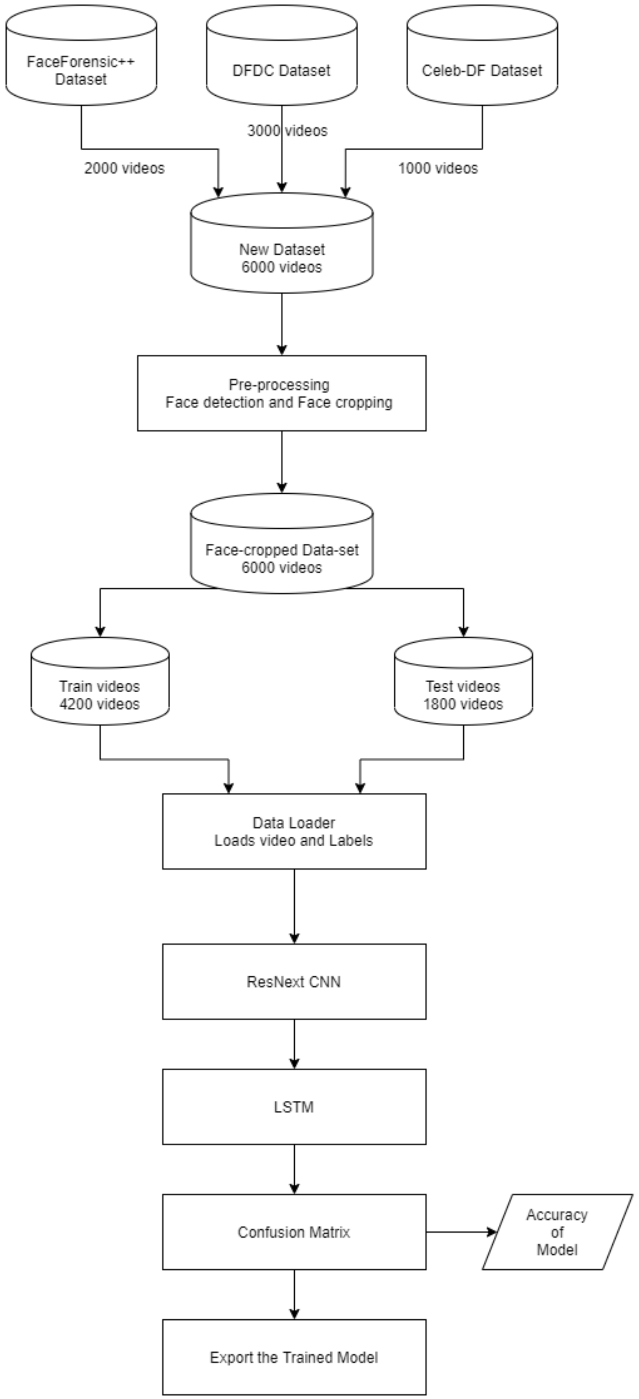
Sequential Layer

1 LSTM layer with 2048 shape input vector and 2048 latent features along with 0.4 chance of dropout and ReLU Activation function





Training Workflow



FaceForensic++ Dataset

DFDC Dataset

Celeb-DF Dataset

2000 videos

3000 videos

1000 videos

New Dataset
6000 videos

Pre-processing
Face detection and Face cropping

Face-cropped Data-set
6000 videos

Train videos
4200 videos

Test videos
1800 videos

Data Loader
Loads video and Labels

ResNext CNN

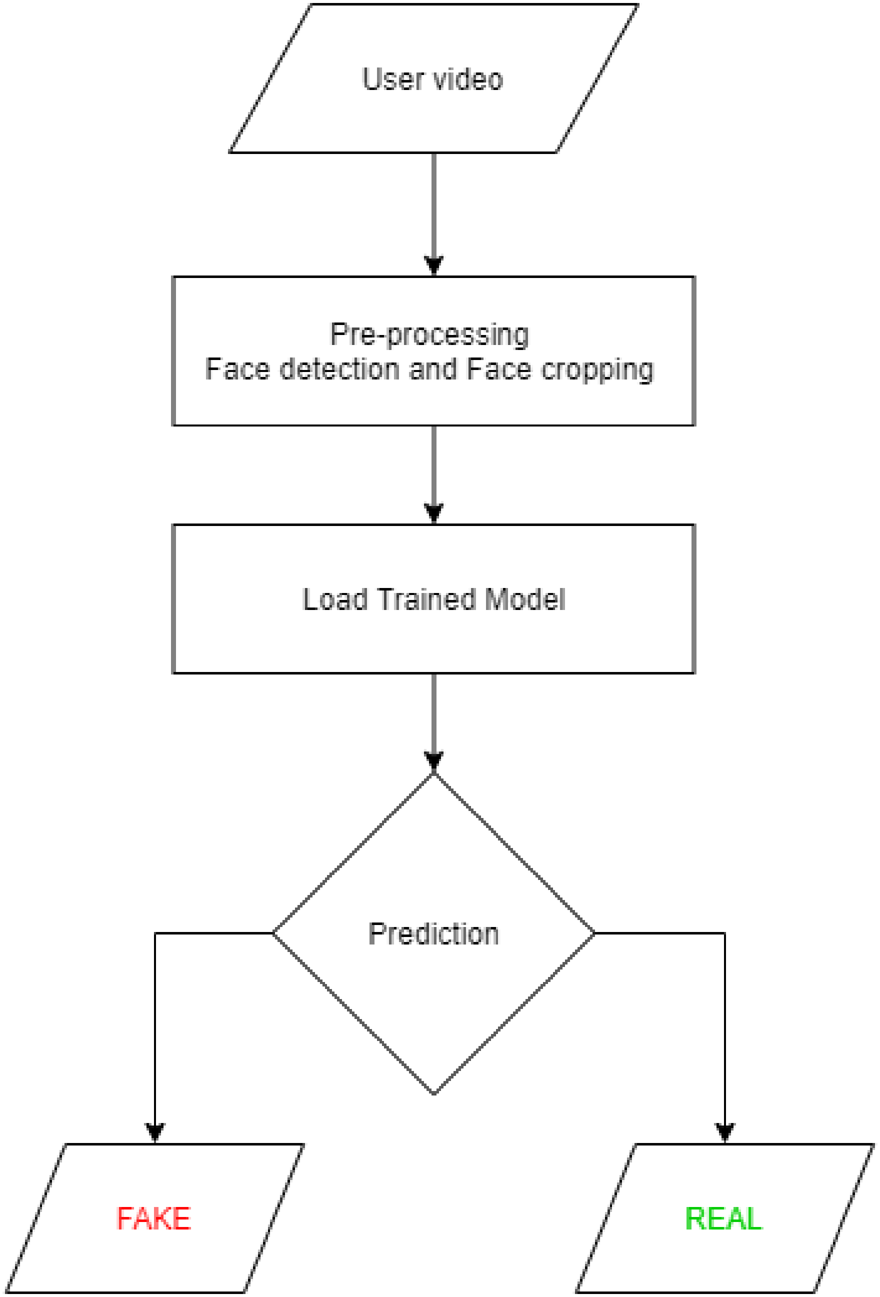
LSTM

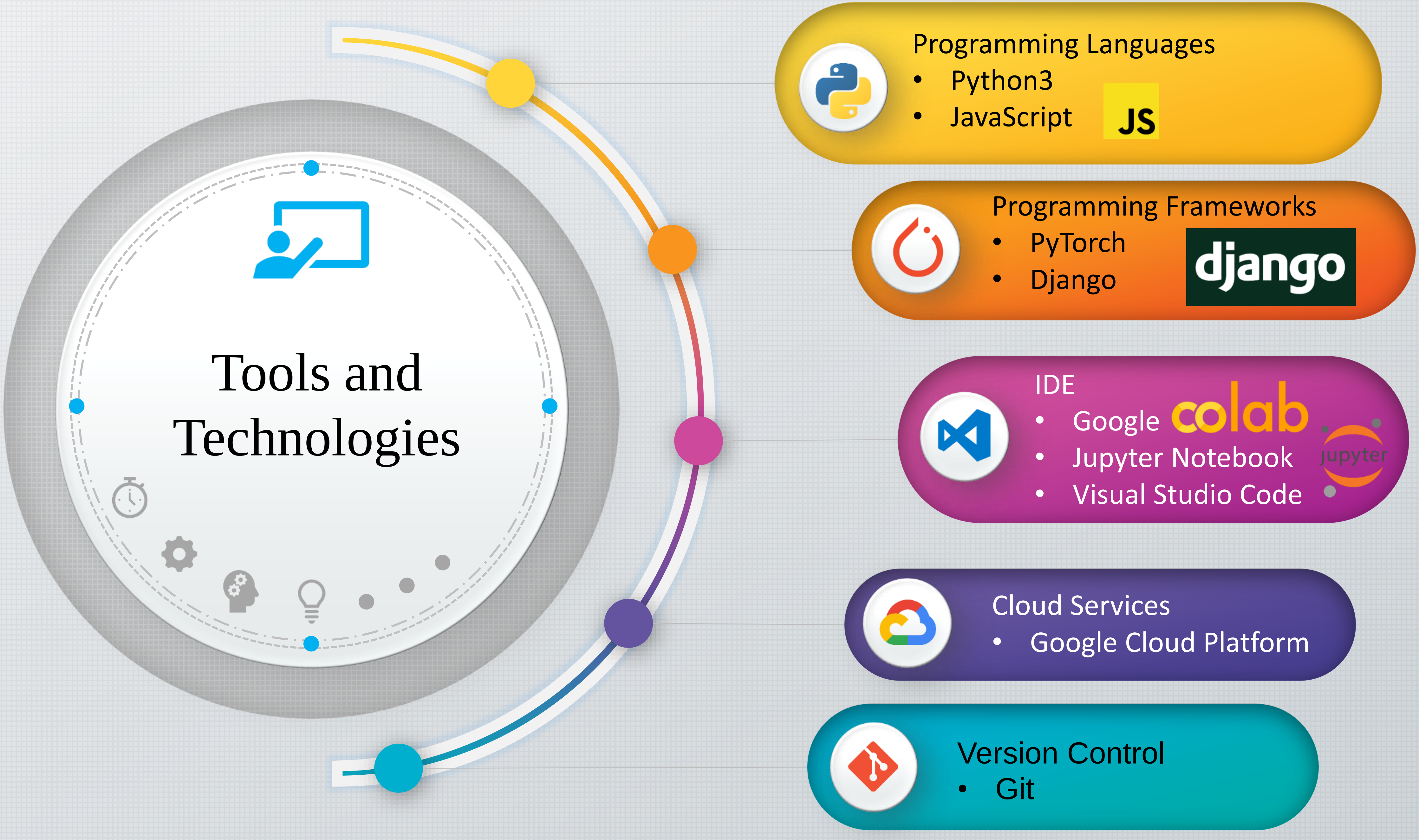
Confusion Matrix

Accuracy of Model

Export the Trained Model

Prediction Workflow





Tools and Technologies

Programming Languages

- Python3
- JavaScript

Programming Frameworks

- PyTorch
- Django

IDE

- Google Colab
- Jupyter Notebook
- Visual Studio Code

Cloud Services

- Google Cloud Platform

Version Control

- Git

Thank you!!