FOOTBALL FLIGHT CURVE PREDICTION

GET THOSE BEAUTIFUL CURVES

Adriel Chen | DSI 23 | Capstone

Problem Statement

A football coach wants to leverage technology to help promote his business and has asked me to come up with something useful that he can use easily.

The type of the flight of the ball depends on many different factors

They include the following and many other factors:

- The angle of approach
- The run-up
- The angle of body parts
- Where the foot strike the ball
- Where the ball is struck
- The kick follow-through

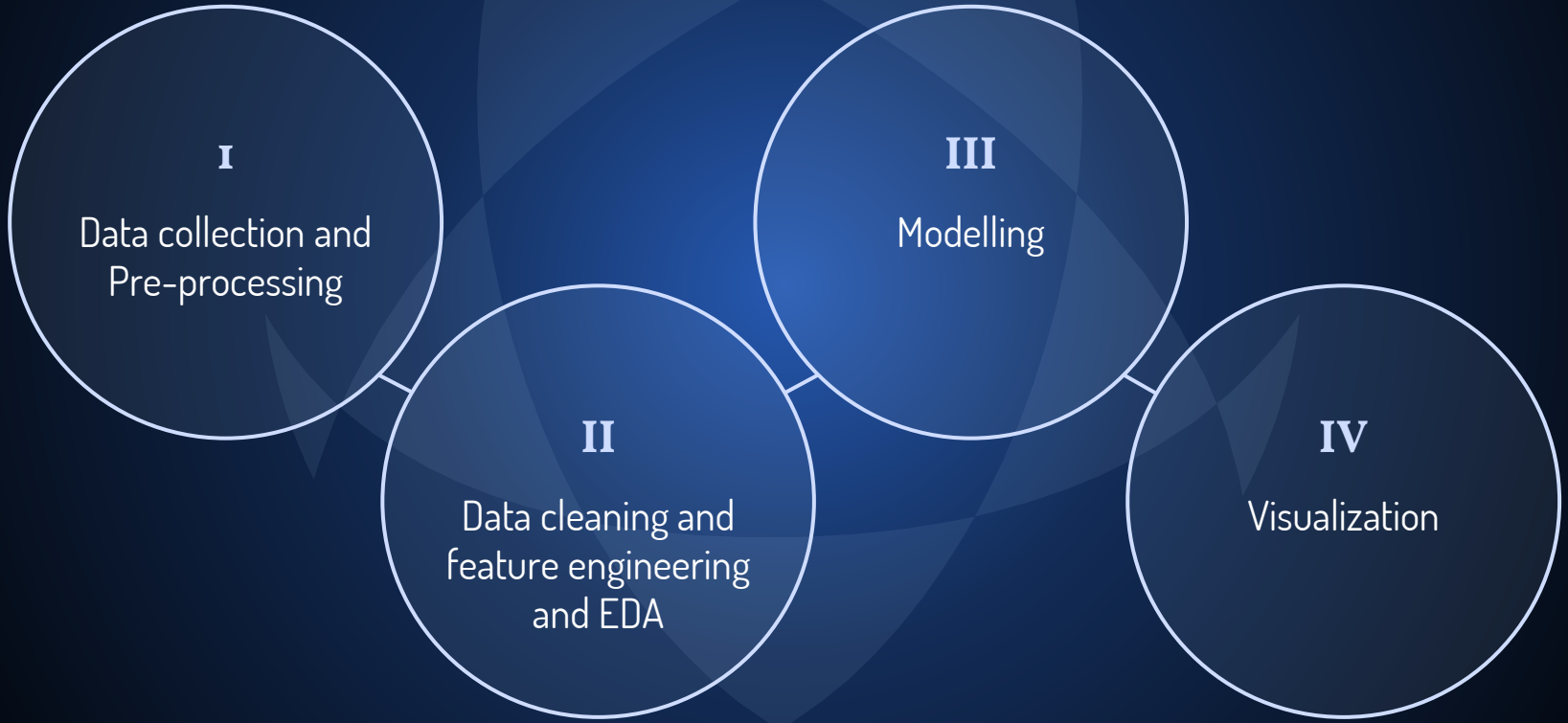
The goal

Different situations call for different kind of football kicks.

The goal is to create a feedback application which returns information on the user on their football kicks to give the everyday footballer more tools to analyze themselves for training purposes.

This project will only focus on curve balls.

APPROACH METHODOLOGY



DATA COLLECTION

Youtube

Collected almost 150 videos of free kicks

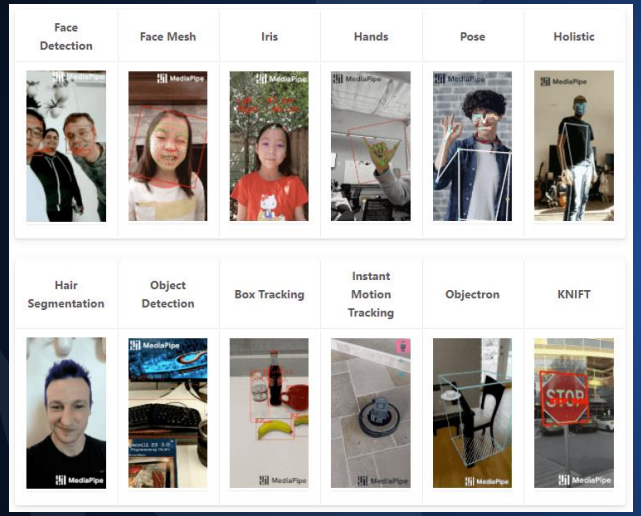
Soccer Event DS

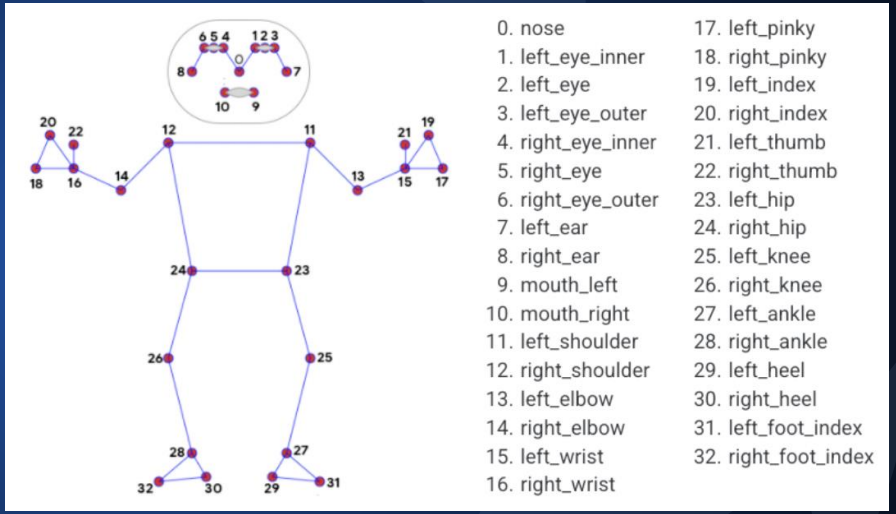
Collected ~30 frames

Mediapipe, OpenCV and Davinci Resolve

DAVINCI
RESOLVE

Mediapipe's Pose Estimation





VIDEO

As there were many people in football matches, mediapipe had trouble focusing on the target

MASKING

Decision was made to gather angles from single frames instead of the whole video.

Masking was done manually in Da Vinci Resolve

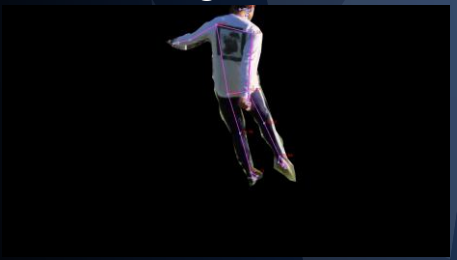
MASKED FRAMES

Frames were stringed together into a video and loaded into notebooks

Data Collection

Curved Video

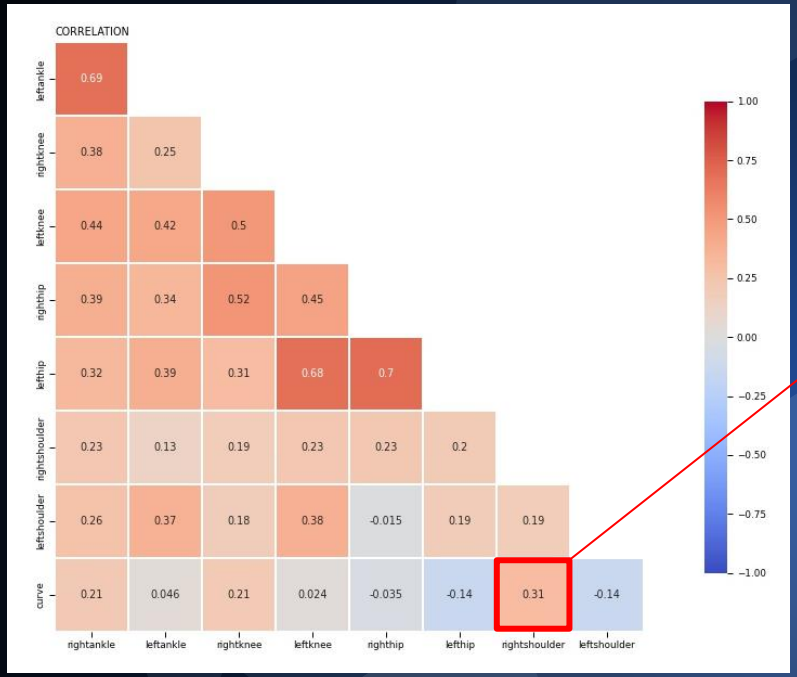
Straight Video



	rightankle	leftankle	rightknee	leftknee	righthip	lefthip	rightshoulder	leftshoulder
0	174.74	177.54	179.82	179.36	176.70	174.20	16.51	19.09
1	176.28	165.74	176.93	176.13	172.89	179.28	25.28	14.28
2	169.27	168.83	174.27	178.06	179.90	172.50	14.30	19.82
3	176.61	154.01	177.29	167.83	162.96	167.97	18.23	6.98
4	174.05	170.13	173.36	177.71	176.88	175.12	13.58	12.62
...	...	...	...	...	...	...	...	...
56	129.30	125.25	148.42	148.43	136.94	149.95	3.78	12.56
57	113.34	128.69	149.08	146.32	137.01	146.76	0.54	11.54
58	86.42	106.52	135.56	163.05	136.04	165.35	3.72	21.80
59	132.82	127.81	167.21	154.73	149.75	144.62	7.61	21.71
60	122.03	117.99	139.77	162.23	135.55	161.85	1.75	15.99

Cleaning & EDA

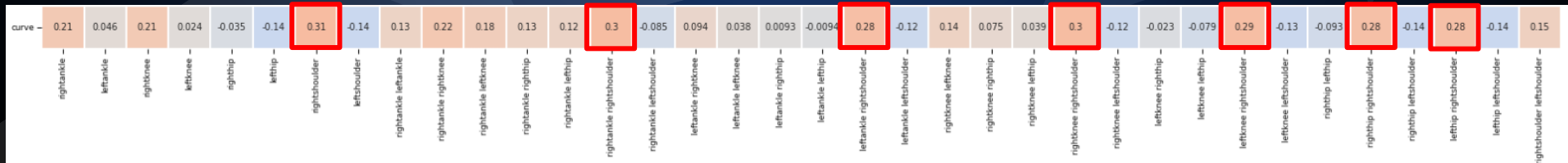
Exploratory Data Analysis



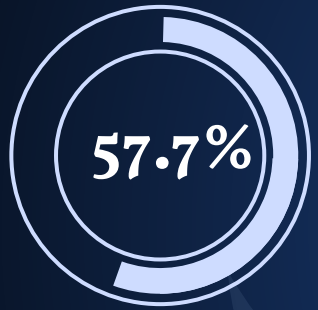
Not much correlation between lower body angles and curve shots

But interestingly, the most correlated feature to a curve ball is the shoulder angle

As expected even after polynomial feature engineering, the highest correlations were features with right shoulder.



Modelling



Baseline score is 57.7%

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
et	Extra Trees Classifier	0.8400	0.9133	0.8700	0.8512	0.8517	0.6680	0.6822	0.0250
rf	Random Forest Classifier	0.8389	0.8997	0.9033	0.8245	0.8568	0.6663	0.6861	0.0290
lightgbm	Light Gradient Boosting Machine	0.8189	0.8898	0.8167	0.8614	0.8229	0.6300	0.6471	0.0130
knn	K Neighbors Classifier	0.7978	0.8871	0.7767	0.8540	0.8014	0.5937	0.6122	0.0040
xgboost	Extreme Gradient Boosting	0.7778	0.8730	0.7933	0.8252	0.7874	0.5407	0.5669	0.0560
qda	Quadratic Discriminant Analysis	0.7678	0.8028	0.9833	0.7189	0.8263	0.5022	0.5514	0.0020
gbc	Gradient Boosting Classifier	0.7578	0.8390	0.8000	0.7769	0.7806	0.5042	0.5220	0.0120
dt	Decision Tree Classifier	0.7478	0.7417	0.7833	0.7721	0.7706	0.4876	0.5016	0.0020
ridge	Ridge Classifier	0.7367	0.0000	0.7033	0.7888	0.7323	0.4687	0.4844	0.0020
lda	Linear Discriminant Analysis	0.7267	0.7035	0.7033	0.7755	0.7262	0.4454	0.4603	0.0020
lr	Logistic Regression	0.7256	0.7142	0.7033	0.7688	0.7226	0.4540	0.4691	0.0230
ada	Ada Boost Classifier	0.6967	0.7972	0.7167	0.7445	0.7219	0.3886	0.4010	0.0120
svm	SVM - Linear Kernel	0.6444	0.0000	0.7633	0.5492	0.6258	0.2355	0.2716	0.0020
nb	Naive Bayes	0.6267	0.6778	0.6100	0.6821	0.6306	0.2663	0.2693	0.0020

Extra trees classifier provided highest accuracy and AUC with similar F1 score to Random Forest

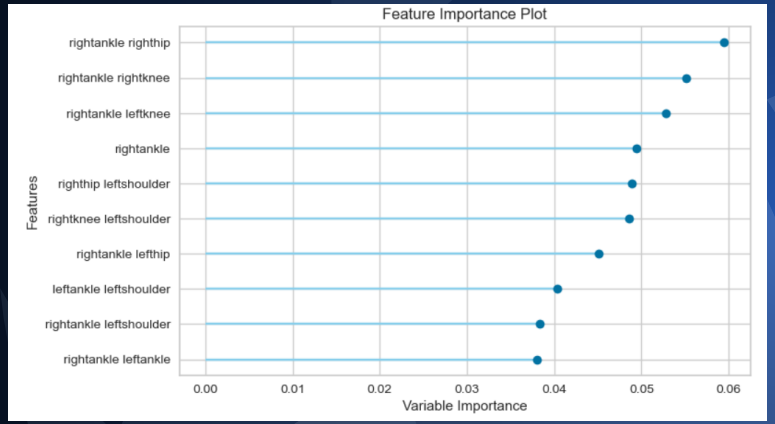
Model Comparison

	Accuracy	AUC	Precision	Recall	F1 Score
Extra Trees	84%	0.9133	0.8512	0.87	0.8517
Bagged Extra Trees	85%	0.926	0.8788	0.8567	0.8627
Boosted Extra Trees	87%	0.9237	0.8679	0.9067	0.8822
Blended (ET, RF LightBGM)	84%	0.9017	0.8621	0.8533	0.8533

Model Comparison

	Accuracy	AUC	Precision	Recall	F1 Score
Bagged Extra Trees (Train)	85%	0.926	0.8788	0.8567	0.8627
Bagged Extra Trees (Test)	86%	0.93	1	0.778	0.875
Boosted Extra Trees (Train)	87%	0.9237	0.8679	0.9067	0.8822
Boosted Extra Trees (Test)	81.4%	0.9363	0.9524	0.74	0.8333

Feature coefficients



Data was manipulated and gathered using only players which struck with the right foot

Top 5 features are the following:

1. Right Ankle x Right Hip
2. Right Ankle x Right Knee
3. Right Ankle x Left Knee
4. Right Ankle
5. Right Hip x Left Shoulder

Visualization

- Relevant limb angles
- Prediction Result
- Ball Speed

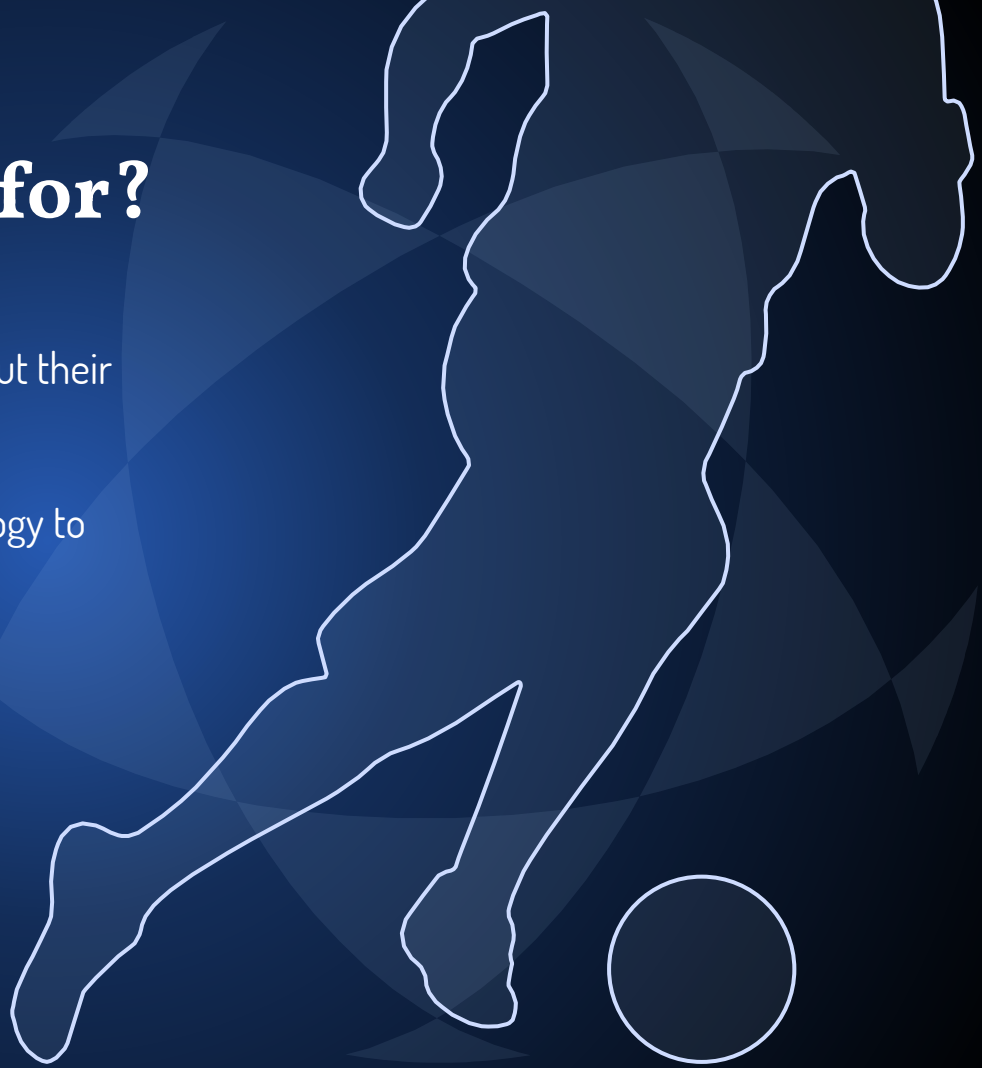
Automatically stores image before release

Automatically stores image before strike

Automatically saves flight path for posterity

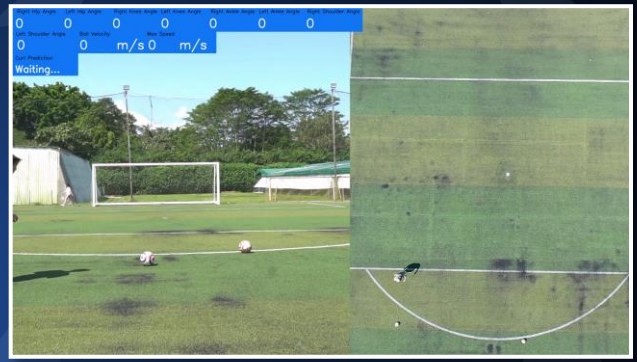
What this app is good for?

1. Social footballers who just want to check out their form, speed or technique
2. Football trainers who want to use technology to market their businesses



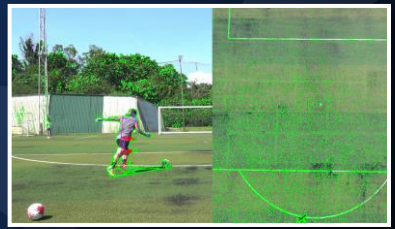
Possible future work

- Skeletal overlay with optimal angles
- Curve angle calculation
- Speed of kick calculation
- Speed of ball from back view
- Run up recording and processing
- Live demo for uploading user videos
- Try to deploy and make a full app



Things for improvement

- Allowing user inputs to set the region of interest for masking detection
- Autodetection of video resolution
- Learn more about computer vision masking for detection accuracy
- Improving pose estimation accuracy by trying with other packages



Special Thanks

Kishan S

Encourager, idea generator

Sharmaine C

Encourager, resource generator

Nicholas R Youtuber with free code on some pose estimation tutorials

Murtaza H Youtuber with free code on OpenCV painting

Sergio C Youtuber with object detection tutorial

THANKS!

QUESTIONS?