# THE MACHINE LEARNING CANVAS Designed for: Deep Learning Course Designed by: Group 3 Date: 20-02-2024 Iteration: 1 .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PREDICTION TASK Instance segmentation on football players in a football match differentiating between teams.  Possible outcomes for the segmentations are players of team A and players of team B and a super category for all other pixels.  For the instances we use rectangular bounding boxes to track where players are in the image. Observations are made in real time. | DECISIONS Predictions are turned into proposed value for the end-user by allowing tracking players, analysing player movement behaviour and player performance based on positional data. Instance segmentation segments players and draws bounding boxes around players. These bounding boxes can be used to track the players. In combination with a depth camera, exact player location can be extracted. Segmentation also allows the analysis of exact player movements like posture, direction, and speed. | VALUE PROPOSITIONThe end-users consist of users like; sports analysts of football teams, betting vendors, and sports websites. The objectives of sports analysts could be determining overall positional tactics, individual player positioning and movements, e.g. is the player free to pass to or not, or are they defending/attacking well. The betting vendors could use these same analyses to determine in-game odds. Analysts have a strong interest in these tactics to analyse games afterwards, whereas betting vendors can monetize this model directly by making more accurate predictions based on player predictions.MONITORING For monitoring user feedback metrics can be used to quantify value creation. Additionally, direct monetization for betting vendors can be measured. Also, statistical tests and drift detection algorithms can be applied to detect model drift. Once this has reached a certain threshold the model should be re-trained. | DATA COLLECTIONThe initial train set is a dataset found on Kaggle at https://www.kaggle.com/datasets/ihelon/football-player-segmentation/data. In the case that continuous updates are necessary, many other labelled data sets are available. It is also possible to extract our own data, however in that case we would need to annotate the images. In this case we could use open-source models like Meta AI’s Segment Anything to swiftly annotate the segmented players. | DATA SOURCES Data sources where we can get raw information on the entities and observed outcomes are basically any professional football match, ideally football matches at the highest professional level due to speed and intensity of the players, and the video capture quality. There are several REST API’s that allow programs to download football match statistics and images. YouTube can also be used as a data source. |
| IMPACT SIMULATIONModels can be deployed through REST API’s or Web Applications. They can be hosted on Google Cloud Platform. In the case that the model is deployed as component of a web app, the user can be provided with the instances and segmentations, as well as pre-computed statistics correlated with specific types of analyses. Gains are that analysts do not need to do any manual labour or having to buy expensive tracking software based on GPS. The costs for incorrect decisions are that analyses might be wrong. Regarding the fairness constraints, the only discriminatory factor in this task is player and team. | MAKING PREDICTIONSBatch size could be used in specific applications but for use cases discussed here we will only consider real-time predictions. Featurization is not necessary as raw images are fed into the neural network. Post-processing can be used to compute statistics like percentage of time attacking or defending. These can be continuously updated in real-time as well. Compute target is a compute instance on Google Cloud Platform or Cloud Run, a serverless solution. | BUILDING MODELSOnly one version of a production model is necessary for this task, it can be run on Cloud Run to improve scalability. Updates would occur when a new batch of training data is collected or when the model has drifted. | FEATURESInput representations fed to the neural network at inference are raw (RGB) images of a currently streaming football match. |

|  |  |  |
| --- | --- | --- |
|  | Version 1.1. Created by Louis Dorard, Ph.D. Licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/). Please keep this mention and the link to [ownml.co](https://www.ownml.co/) when sharing. | [**OWNML.CO**](https://www.ownml.co/) |