# THE MACHINE LEARNING CANVAS Designed for: Datascientest Project Designed by: Simon, Hakan Date: 14.05.2024 Iteration: 0 .

| PREDICTION TASKType of task? Entity on which predictions are made? Possible outcomes? Wait time before observation?Classify the input heartbeat pattern and assign it to either 5 classes or a binary distinction in healthy or non-healthy heartbeat pattern.The heartbeats are delivered in a standardized and normalized format in csv-files.There is no wait time before observation, because the heartbeat patterns are delivered in “real time”. | DECISIONSHow are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that. The decision maker would be the medical staff. We do not know how they would handle the output of our model, but would advise the following: Use the models to accelerate the decision making when vast amounts of data have to be analyzed at once. Define a certain threshold (we do not do this, because it is unethical) on which patients are double-checked for heartbeat failure or possible failures and or related heartbeat failures. | VALUE PROPOSITIONWho is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces. End-users are employees in the medical field. Their objective is to maximize the accuracy of heartbeat pattern detection and to automate it as far as possible to save time and human resources.  They benefit from the ML-System because the System can make a first estimation on what the observed heartbeat pattern presents. If a certain threshold is given, a certain number of patients has not to be checked manually by the medical staff.  Possible workflows / interfaces:  The Model has to be integrated in the measuring device (sampler of electrode signals) and continuously detect the given signals. This can be done via microprocessors that contain the trained models.  If the model detects heartbeat patterns over a certain threshold, they give alarm to the medical staff, then a double check will be performed. | DATA COLLECTIONStrategy for initial train set & continuous update. Mention collection rate, holdout on production entities, cost/constraints to observe outcomes. The initial dataset is gathered from Kaggle: https://www.kaggle.com/datasets/shayanfazeli/heartbeat  Proposed Workflow outside of the microcontroller (MLflow): All Data is gathered together, along with the feedback of the medical staff when heartbeat failure has been detected. A routine is implemented to check the current version of the model against the new data and feedback from the medical staff. The model is then retrained, maybe according to a certain number of wrong predictions or after a fixed amount of newly gathered data. → We cannot simulate this continuous update so far, only with oversampling some simulation of it would be possible. | DATA SOURCESWhere can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc. Only oversampling for generation of “new data” is possible, because the labeling of new data would require expert knowledge, which we do not have.  Raw data is available easily, but needs to be labeled to be useful as training data. The labeling is time and cost intensive and can only be done by medical experts.  There is no online library which is continuously updated with labeled data. Also non-labeled data is not available online.  Best practice would be to investigate the research papers on this topic to gather new data. |
| --- | --- | --- | --- | --- |
| IMPACT SIMULATIONCan models be deployed? Which test data to assess performance? Cost/gain values for (in)correct decisions? [Fairness constraint](https://developers.google.com/machine-learning/glossary#fairness-constraint)? Models can be deployed directly to microcontrollers or to an API on which the microcontrollers could have access to.  Train and Test data was used from the original dataset as mentioned in chapter “data collection / data sources”.  If the model predicts false negatives, people could suffer from undiagnosed heartbeat failures.  We don´t have implemented a fairness constraint yet as this would require extensive collaboration with the medical stuff that would define such constraints / thresholds that manipulate our raw output. | MAKING PREDICTIONSWhen do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target? The model can be used for both prediction tasks:  The model itself always works with a single heartbeat which is delivered in form of a np.array.  If real time prediction is necessary, a small workaround could be implemented to deliver the data not in csv-files but rather directly in dataframes. If batch analysis is necessary, the models work directly from the get-go with csv-files (wrap around function to feed them is implemented also). Realtime prediction would be usecase for medical staff at work, batch would be training the model for better accuracy. The model training takes about 2-5 hours, depending on computing power and size of training set. Prediction time on single heartbeat is negligible. No Post-Processing is done since the desired heartbeat classifications are the output. |  | BUILDING MODELSHow many prod models are needed? When would we update? Time available for this (including featurization and analysis)? We have done the Datascience Project on this topic and produced 8 models. Only the best model should be deployed and continuously updated.  Updating is done on thresholds that the medical staff would define.  We can update the model on a dedicated machine that is not running the realtime production, so no disturbances occur. Then one after another, new instances of the updated model could be deployed. | FEATURESInput representations available at prediction time, extracted from raw data sources. The available data was extensively preprocessed from the original authors.  Single heartbeats were extracted through a specific cross-correlation method, which was done on both train and test set. We do not have access to this function and can therefore not generate the exact same data as we used.  The procedure (but not the function) for single heartbeat extraction can be found in the original paper: <https://arxiv.org/pdf/1805.00794> |
|  | MONITORINGMetrics to quantify value creation and measure the ML system’s impact in production (on end-users and business)? accuracy and recall. We can give those metrics to the medical staff and explain what they mean. Then they have to decide how the thresholds for giving an double-check-alarm and for re-deploying the model have to be defined. |  |  |  |

|  | 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/) |
| --- | --- | --- |