

Use of machine-learning algorithms to predict urine culture results in cats

Materials and methods

Feature preprocessing

Patient data, including sex, age, color, clarity, USG, and ph, were utilized as features. Numeric features: protein, glucose, ketones, hemoprotein, bilirubin, WBC, and RBC were accounted for to target positive culture. Ignore_features (visit number; BW) were excluded from training as its values contain NaN values and strings. A positive culture was encoded using the one-hot encoding (0 and 1). Cat sex was splitted into 4 different groups.

Machine-learning model training

Two Feline UTI prediction models were trained. The ExtraTreesClassifier (ET) model was trained with all patient features excluding Sed Bac Quantity (SBQ) and Sed Bac Type (SBT). The second model, ExtraTreesClassifier (ET) trained with excluded features SBQ; SBT. The first model excluded SBQ and SBT due to limited resources to obtain data from the veterinary lab. The second model was trained to analyze AUC performance by just training the excluded data.

	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Fold							
0	0.8662	0.7313	0.3846	0.7692	0.5128	0.4451	0.4811
1	0.8944	0.8335	0.5577	0.8056	0.6591	0.5990	0.6132
2	0.9085	0.8433	0.6346	0.8250	0.7174	0.6639	0.6720
3	0.8768	0.8497	0.5192	0.7297	0.6067	0.5361	0.5470
4	0.8662	0.7933	0.4615	0.7059	0.5581	0.4833	0.4985
5	0.8627	0.8236	0.4423	0.6970	0.5412	0.4651	0.4818
6	0.8905	0.8100	0.5490	0.7778	0.6437	0.5812	0.5935
7	0.8834	0.7787	0.5192	0.7714	0.6207	0.5549	0.5701
8	0.8587	0.7327	0.3654	0.7308	0.4872	0.4156	0.4493
9	0.8905	0.8524	0.5769	0.7692	0.6593	0.5957	0.6044
Mean	0.8798	0.8048	0.5011	0.7582	0.6006	0.5340	0.5511
Std	0.0155	0.0429	0.0817	0.0394	0.0698	0.0754	0.0681
First model (Data_limited)							

	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Fold							
0	0.8908	0.7987	0.5192	0.8182	0.6353	0.5749	0.5954
1	0.9014	0.8077	0.6346	0.7857	0.7021	0.6439	0.6491
2	0.8873	0.8375	0.5769	0.7500	0.6522	0.5863	0.5935
3	0.8838	0.8003	0.5769	0.7317	0.6452	0.5768	0.5827
4	0.8662	0.8278	0.5577	0.6591	0.6042	0.5243	0.5270
5	0.8944	0.8579	0.6346	0.7500	0.6875	0.6245	0.6276
6	0.8410	0.7543	0.4706	0.5714	0.5161	0.4221	0.4249
7	0.8657	0.7472	0.4231	0.7333	0.5366	0.4646	0.4887
8	0.8481	0.7878	0.4808	0.6098	0.5376	0.4482	0.4528
9	0.8516	0.7454	0.4038	0.6562	0.5000	0.4186	0.4356
Mean	0.8730	0.7965	0.5278	0.7065	0.6017	0.5284	0.5377
Std	0.0203	0.0367	0.0777	0.0750	0.0700	0.0802	0.0786
Second model (SBQ_SBT)							

Figure 1. Data table illustrating AUC performances between two different models.

The ET model (Fig.1) utilized the training set data, allowing us to determine how numerical features affect the prediction performance for Feline UTI. The model trained using a repeated 10-folds for each of the 10 candidates, totaling 100 times to produce the average confusion matrix to tune the hyperparameters, using the built-in function:

Tuned_et = tune_model(et) (Python3; ipykernel).

The boundary plot (Fig. 2) visualizes how plots spread for trained set data. Through visualization, we can analyze that all trained data sets were spread out for the first model (Data_limited), and the decision boundary will be in the negative linear slope form. While the second model (SBQ, SBT) reflects skewed trained data sets on the right-hand side of the plot, it will include the decision boundary in the form of a curve hyper-plane to fit the dimension of the space.

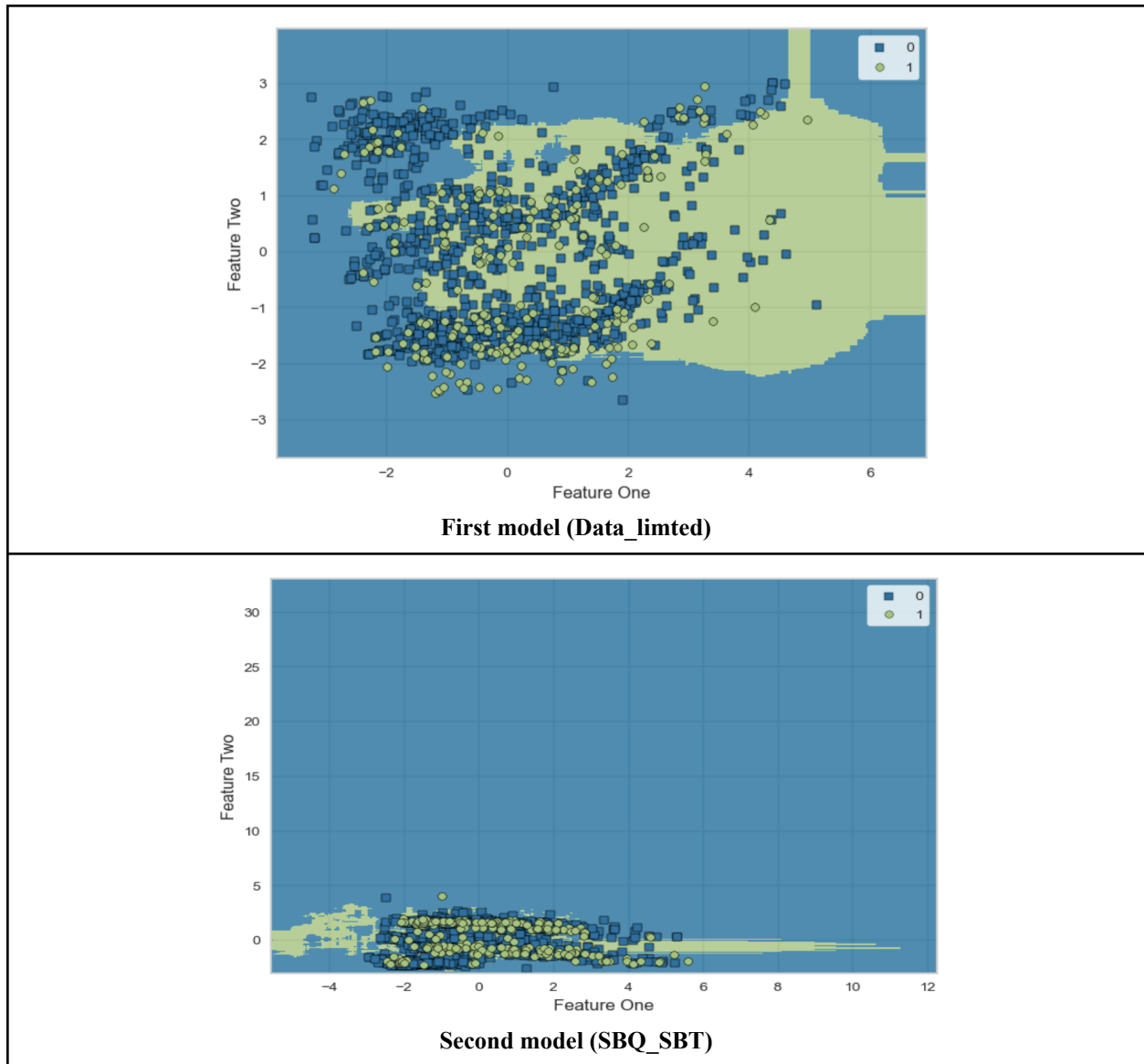


Figure 2. Data table illustrating AUC performances between two different models.

Machine-learning model performance evaluation

The test set included [types of cats]. The hyperparameters optimized on the training set utilized during ET (with all features excluding sed features) and ET (with only sed features) model testing. Model prediction results reported sensitivity and specificity. Predictions made utilized as a binary classifier, a receiver operating characteristic (ROC) plot was generated, and the area under the curve (AUC) was calculated.

Results

During the study period, 6,732 unique cats were evaluated. Of these cats, 5,850 were tested for Feline UTI prediction: 4,634 in the training set, and 1,216 in the test set. Under the category of “Positive culture”, there were a total of 4,052. Between collected data, 3,310 of the training sets were under 0 while 742 remaining data sets were set

to 1. The (Fig. 3) displays the summary of collected statistics for the training and test set. The data was measured using the built-in function:

`.describe()` (Python3; ipykernel). 0 - neg / 1 - pos

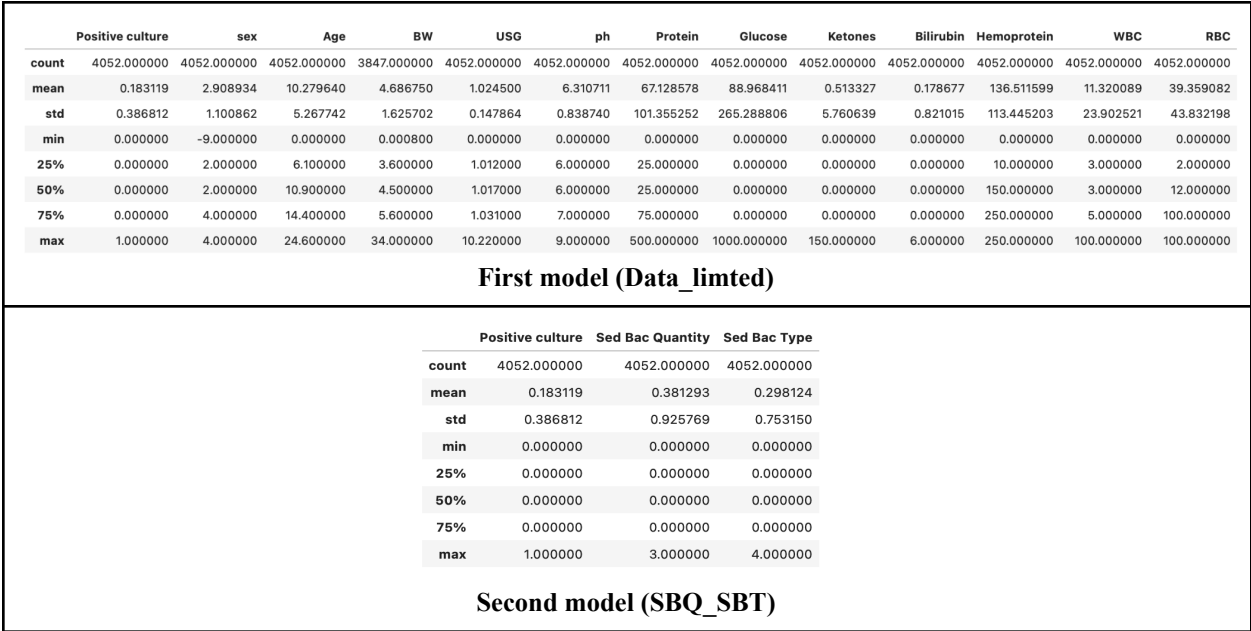


Figure 3. Data table describing the summary of statistics.

Machine-learning model performance

The ET model (Fig. 1) displays an AUC performance of 80.48% (first model) compared to the second ET model 79.65%. The ET first model showed slightly higher accuracy than the second model by 0.68%. The first ET model performed better on AUC, Accuracy than the second model of ET.