Predictive Maintenance of Armoured Vehicles using Machine Learning Approaches

Prajit Sengupta, Anant Mehta and Prashant Singh Rana

Abstract

The proposed ensemble model approach involves the use of various models such as Light Gradient Boosting, Random Forest, Decision Tree, Extra Tree Classifier and Gradient Boosting to accurately predict the maintenance requirements of the vehicles.

Dataset

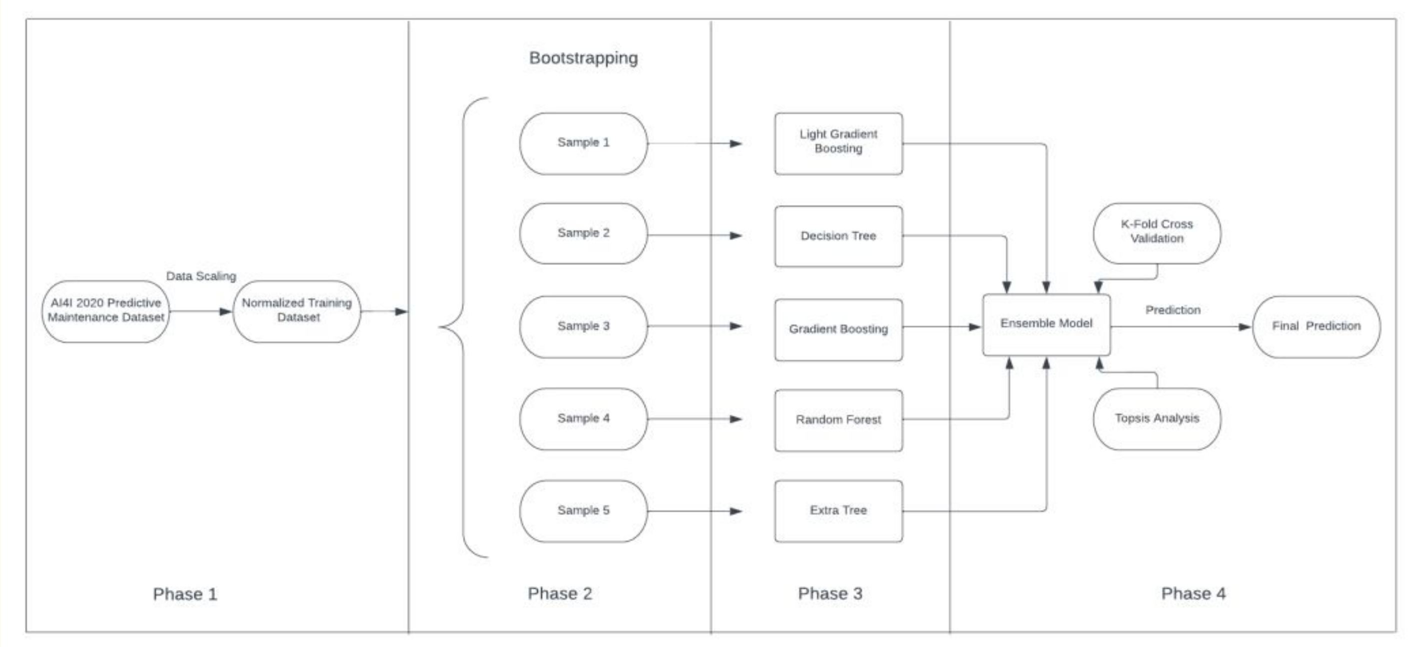
Synthetic dataset that aims to reflect real predictive maintenance data!

- AI4I 2020 Predictive Maintenance Dataset(UCI) Repository
- 10,000 rows of data, with six features stored in columns.
- Product Type, Air Temp, Process Temp, Rotational Speed, Torque, Tool Wear[1]

TABLE I
FEATURE TABLE

Machine Number	F1	F2 [K]	F3 [K]	F4 [rpm]	F5 [Nm]	F6 [min]	Machine failure
0	2	0.304347826	0.358024691	0.222933644	0.535714286	0	0
1	1	0.315217391	0.37037037	0.139697322	0.583791209	0.011857708	0
2	1	0.304347826	0.345679012	0.192083818	0.626373626	0.019762846	0
3	1	0.315217391	0.358024691	0.154249127	0.490384615	0.027667984	0
4	1	0.315217391	0.37037037	0.139697322	0.497252747	0.035573123	0
5	2	0.304347826	0.358024691	0.149592549	0.523351648	0.043478261	0
6	1	0.304347826	0.358024691	0.227008149	0.53021978	0.055335968	0
7	1	0.304347826	0.358024691	0.208963912	0.5	0.063241107	0
8	2	0.326086957	0.37037037	0.290454016	0.340659341	0.071146245	0
9	2	0.347826087	0.407407407	0.333527357	0.332417582	0.083003953	0
.		.		.		.	
.		.		.		.	
1353	2	0.358695652	0.543209877	0.160651921	0.608516484	0.154150198	0
1354	1	0.358695652	0.530864198	0.096623981	0.607142857	0.166007905	0
1355	2	0.358695652	0.530864198	0.298603027	0.43543956	0.173913043	0
1356	1	0.358695652	0.518518519	0.247380675	0.461538462	0.185770751	0
1357	1	0.358695652	0.518518519	0.170547148	0.575549451	0.193675889	0
1358	2	0.369565217	0.543209877	0.305587893	0.385989011	0.201581028	0
1359	1	0.369565217	0.555555556	0.137369034	0.638736264	0.213438735	0
1360	0	0.380434783	0.555555556	0.313736903	0.326923077	0.221343874	0
.		.		.		.	
.		.		.		.	
7196	0	0.52173913	0.543209877	0.150174622	0.534340659	0.252964427	0
7197	1	0.510869565	0.530864198	0.185098952	0.521978022	0.272727273	0
7198	2	0.510869565	0.518518519	0.210710128	0.421703297	0.280632411	0
7199	1	0.510869565	0.518518519	0.197322468	0.516483516	0.292490119	0
7200	1	0.510869565	0.518518519	0.256111758	0.406593407	0.300395257	0
7201	1	0.510869565	0.50617284	0.178114086	0.505494505	0.308300395	0
7202	1	0.510869565	0.49382716	0.091967404	0.798076923	0.316205534	0
.		.		.		.	
.		.		.		.	
9992	1	0.380434783	0.333333333	0.183934808	0.486263736	0.031620553	0
9993	1	0.380434783	0.333333333	0.135622817	0.597527473	0.039525692	0
9994	1	0.380434783	0.320987654	0.271245634	0.331043956	0.04743083	0
9995	2	0.380434783	0.333333333	0.253783469	0.353021978	0.055335968	0
9996	0	0.391304348	0.333333333	0.27008149	0.384615385	0.067193676	0
9997	2	0.402173913	0.358024691	0.277648428	0.406593407	0.086956522	0
9998	0	0.402173913	0.37037037	0.139697322	0.614010989	0.098814229	0
9999	2	0.402173913	0.37037037	0.193247963	0.5	0.118577075	0

Proposed Model Workflow



Ensemble Model Using Bootstrap Aggregation

Proposed Model

1. Data creation, normalization and splitting.

2. Trying various models on the dataset and finding out the best hyperparameters for each. Bootstrap samples created.

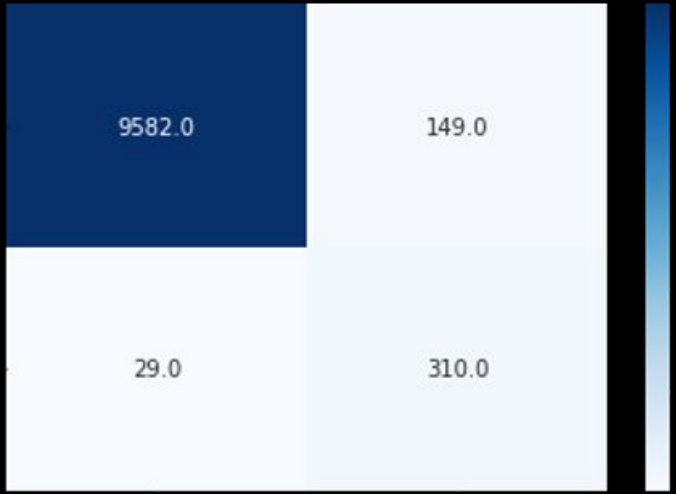
3. TOPSIS[2] used to find the best models.

4. Ranking the models and validating the ensemble using k-fold cross validation[3].

   Best scores were obtained by :

   Light Gradient Boosting, Decision Tree, Gradient Boosting, Random Forest and Extra Tree

Evaluation Metrics



Precision = TP/(TP + FP)

*F 1 = 2 * (pre.* recall)/(pre. + recall)*

Area Under Curve

Recall = TP/(TP + FN)

Kappa = (OA − EA)/(1 − EA)

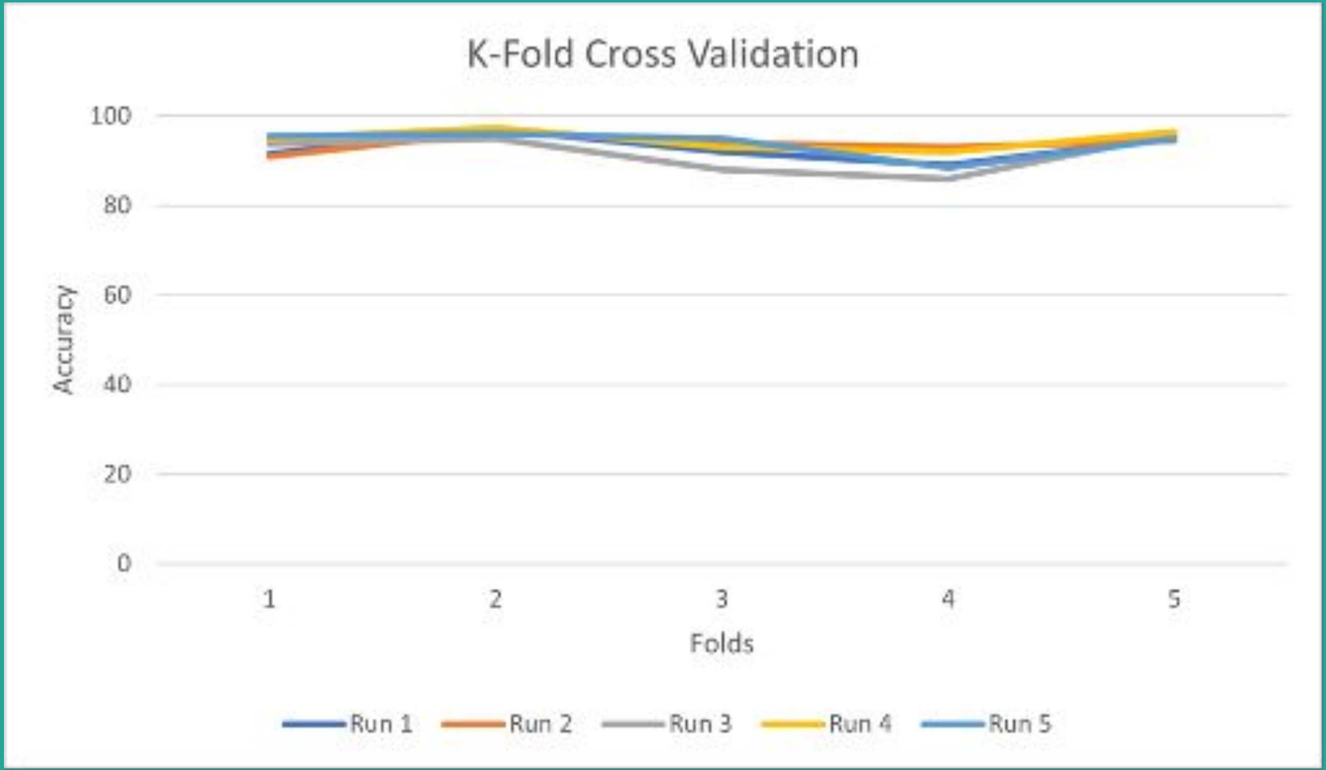
Accuracy = (TP +TN)/(TP +TN +FP +FN)

Matthews correlation coefficient (MCC)

Scores Achieved

- Accuracy : 0.9892
- AUC : 0.977
- Recall : 0.9970
- Precision: 0.9918
- F1 Score : 0.9944
- Kappa : 0.8013
- MCC : 0.8482

Cross Validation Curve (5 Epochs)



Analysis with Different Statistical Models

TABLE III
EVALUATION PARAMETERS

Abbreviation	Model	Accuracy	AUC	Recall	Precision	F1	Kappa	MCC	TT(Sec)
lightgbm	Light Gradient Boosting Machine	0.9847	0.9768	0.6215	0.8748	0.7208	0.7133	0.7273	0.275
dt	Decision Tree Classifier	0.981	0.8539	0.718	0.7103	0.7095	0.6997	0.7021	0.043
et	Extra Trees Classifier	0.9777	0.9602	0.3344	0.9456	0.4906	0.4819	0.5526	0.47
svm	SVM - Linear Kernel	0.9676	0	0	0	0	0	0	0.028
ridge	Ridge Classifier	0.9676	0	0	0	0	0	0	0.019
knn	K Neighbors Classifier	0.9741	0.8194	0.268	0.8194	0.3952	0.3857	0.4527	0.099
ada	Ada Boost Classifier	0.9727	0.9548	0.401	0.6216	0.4834	0.4702	0.4841	0.16
qda	Quadratic Discriminant Analysis	0.8319	0.807	0.4547	0.2939	0.2889	0.2571	0.2729	0.015
gbc	Gradient Boosting Classifier	0.9821	0.9734	0.5911	0.8102	0.6796	0.6707	0.6815	0.434
rf	Random Forest Classifier	0.9817	0.9654	0.4978	0.9018	0.6351	0.6266	0.6589	0.515
lda	Linear Discriminant Analysis	0.9687	0.8748	0.3573	0.5413	0.4241	0.4089	0.4213	0.017
lr	Logistic Regression	0.968	0.8472	0.0132	0.2	0.0247	0.024	0.05	0.389
dummy	Dummy Classifier	0.9676	0.5	0	0	0	0	0	0.013
nb	Naive Bayes	0.961	0.8696	0.2241	0.3451	0.27	0.251	0.258	0.021
em	**Proposed Ensemble Model**	**0.9892**	**0.977**	**0.9970**	**0.9918**	**0.9944**	**0.8013**	**0.8482**	**0.268**

Incorporating additional data sources, such as weather conditions, terrain data, and operator behaviour.

Use of real-time data from sensors and the incorporation of advanced analytics.

Use of Deep Learning in this domain.

Developing new Machine Learning or Mathematical Model for better analysis.

Future Scope

References

[1] Andrea Torcianti and Stephan Matzka. Explainable artificial intelligence for predictive maintenance applications using a local surrogate model. In *2021 4th International Conference on Artificial Intelligence for Industries (AI4I)*, pages 86–88. IEEE, 2021.

[2] Majid Behzadian, S Khanmohammadi Otaghsara, Morteza Yazdani, and Joshua Ignatius. A state-of the-art survey of topsis applications. *Expert Systems with applications*, 39(17):13051–13069, 2012.

[3] Michael W Browne. Cross-validation methods. *Journal of mathematical psychology*, 44(1):108–132, 2000.