

**Technical University of Applied Sciences Würzburg-Schweinfurt  
(THWS)**

Faculty of Computer Science and Business Information Systems

## **Master Thesis**

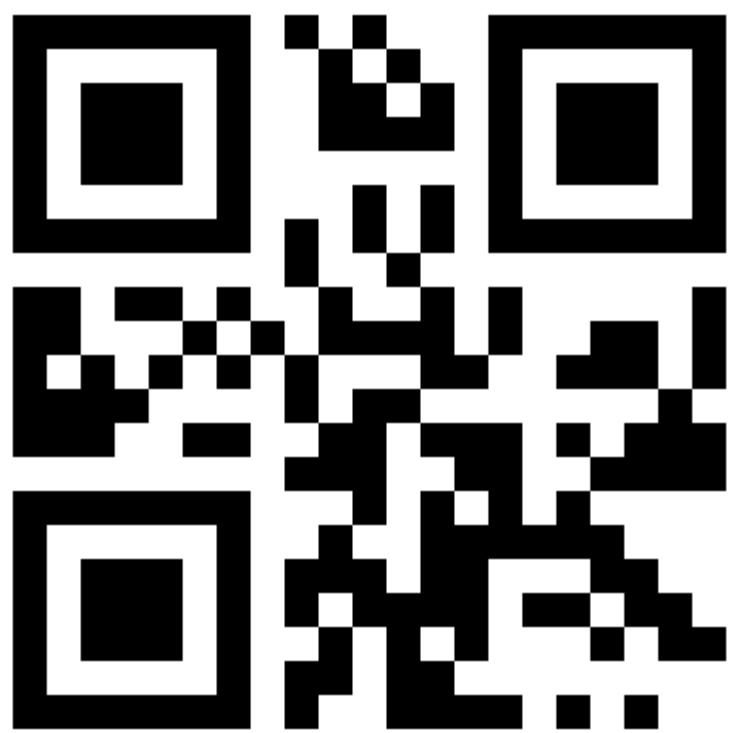
# **Electric Motor Modelling via Graph Neural Networks**

Submitted to the Technical University of Applied Sciences  
Würzburg-Schweinfurt in the Faculty of Computer Science and Business  
Information Systems to complete a course of studies in Master of Artificial  
Intelligence

Lilly Abraham  
K64889

To be Submitted on: 11.12.2024

Initial examiner: Prof. Dr. Magda Gregorova  
Secondary examiner: Prof. Gracia Herranz Mercedes



# Abstract

The thesis explores an approach to predict Key Performance Indicators(KPI)s of topology invariant Interior Permanent Magnet Synchronous Magnets(IPSM) Electric Motors by transforming its geometric, physical and simulation parameters into a graph representation.

The KPIs to be predicted are plots on Efficiency grid(3D) and Torque curve(2D).

We aim to first parameterize the EM design such that it is feasible to convert into a graph representation. Next, we would create a Graph with relevant attributes and design a Graph Neural Network(GNN)with the graph as input and the plots in the format of vectors as target values.

Additionally we may also need to customize the loss function in a way that would smoothen out the plot curves of the prediction values.

Then, we would evaluate the predictions with the test target values by experimenting with various hyper-parameter tuning settings and as a baseline with an Multi Layer Perceptron(MLP) model of the parameters in tabular form.

Finally we will enable the KPI's plot visualisation in a manner presentable to the client Valeo(Automaker Company).

Not necessary remove i suppose.... The aim of the Master Thesis is to train a neural network to learn the parameters of Electric Motors and thus be able to predict its KPIs. The KPIs are 2D and 3D plots on Torque(Mgrenz) curve(Mgrenz) and Efficiency grid(ETA). Other KPIs can be calculated from these two KPIs. For instance the Vibration Costs are inversely proportional to the Efficieny values predicted.

# Acknowledgement

I would like to thank my supervisor Prof. Dr. Magda Gregorova for her guidance and support throughout the course of this thesis and Valeo for providing the data. Special thanks to Mr Daniel and Leo for sharing valuable insights of the data from an electromechanical standpoint.

# Contents

<b>Abstract</b>	<b>2</b>
<b>Acknowledgement</b>	<b>3</b>
<b>Contents</b>	<b>4</b>
<b>Abbreviations</b>	<b>6</b>
<b>Introduction</b>	<b>7</b>
Problem Statement . . . . .	8
Thesis Structure . . . . .	8
<b>Literature Review</b>	<b>9</b>
<b>Dataset</b>	<b>10</b>
Data Preprocessing for Multi Linear Perceptron (MLP) . . . . .	12
Scaling . . . . .	16
Dataset splitting . . . . .	16
<b>Graph Modelling</b>	<b>17</b>
Data Preprocessing for Graph Neural Network (GNN) . . . . .	17
Heterogeneous GNN Model . . . . .	18
<b>Modelling &amp; Evaluation</b>	<b>20</b>
MLP Model . . . . .	20
Loss Functions . . . . .	20
Optimizer . . . . .	22
Evaluation Metrics . . . . .	22
Post Processing . . . . .	23
<b>Experiments and Results</b>	<b>24</b>
Experiments with MLP . . . . .	24
Results with MLP . . . . .	26
Results with Baseline . . . . .	31
Ablation Studies . . . . .	34
<b>Conclusion</b>	<b>36</b>
Future Improvements . . . . .	36

<b>List of Figures</b>	<b>37</b>
<b>List of Tables</b>	<b>38</b>
<b>Bibliography</b>	<b>39</b>
<b>Declaration on oath</b>	<b>40</b>
<b>Consent to Plagiarism Check</b>	<b>41</b>

# Abbreviations

<b>GNN</b>	Graph Neural Network
<b>MLP</b>	Multi Linear Perceptron
<b>KPI</b>	Key Performance Indicator
<b>EM</b>	Electric Motor
<b>FEM</b>	Finite Element Method
<b>CNN</b>	Convolution Neural Network
<b>2D</b>	2 Dimension
<b>3D</b>	3 Dimension
<b>Wandb</b>	Weights & Biases
<b>MSE</b>	Mean Squared Error

# Introduction

In the design of electric motors, vast amounts of data are generated to determine which design of an Electric Motor (EM) fits best to Key Performance Indicator (KPI)s.

KPIs of an Electric Motor are essential to judge the performance of the motor before it is manufactured. Traditionally these KPIs are inferred from a description of an EM design via a Finite Element Method (FEM) approximating the solutions of the Maxwell's equations. This process, though well established in the EM design, is very time consuming and does not allow for high-throughput engine design optimization.

The actual engine data of Valeo is used here as the dataset comprising of multiple variant designs of the Double-V topology.

The 3 motor topologies manufactured by Valeo are as below:

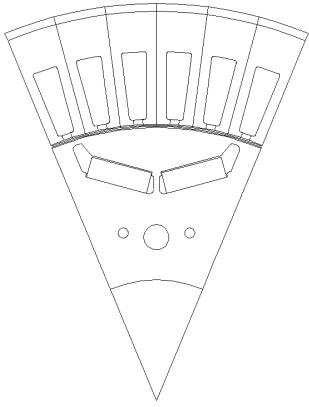


Figure 1: V1 Magnet  
(Source:Valeo)

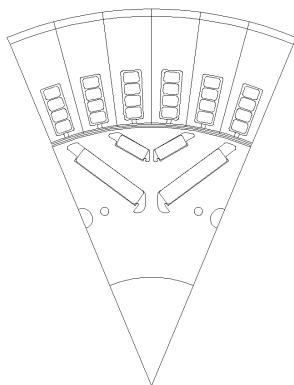


Figure 2: V2 Magnet  
(Source:Valeo)

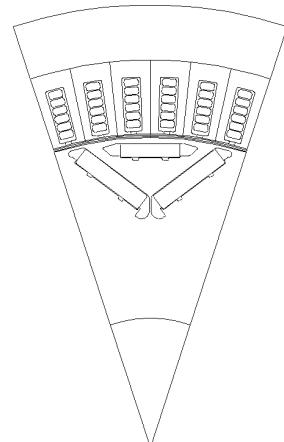


Figure 3: Nabla Magnet  
(Source:Valeo)

This master thesis explores a way to do surrogate modelling of the current process as is highlighted in Figure 4 by making use of GNN or MLP for the modelling of electrical engine designs described parameterically.

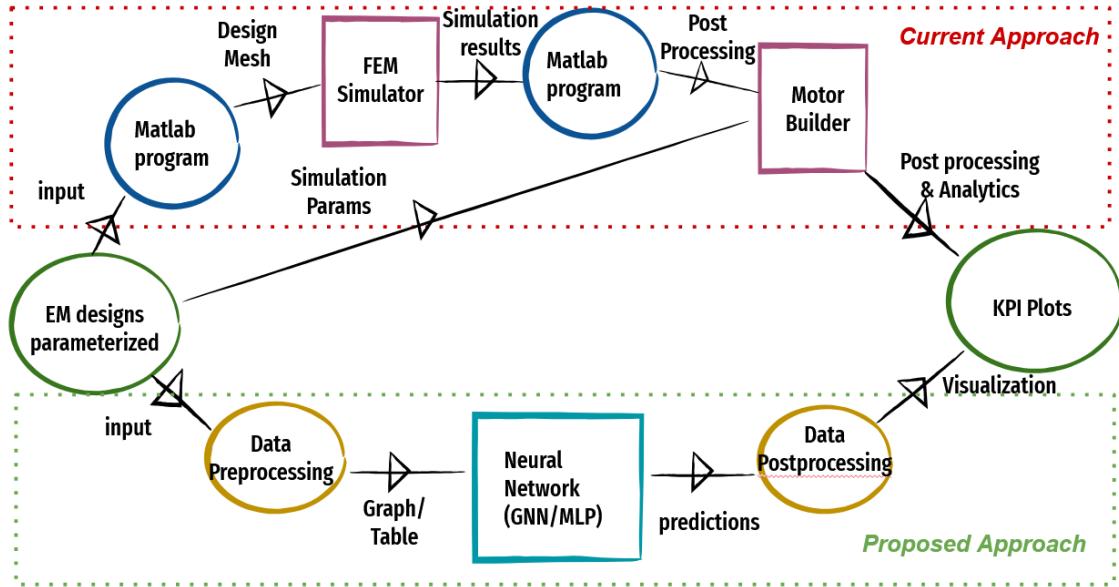


Figure 4: EM Design Flowchart

## Problem Statement/ Research Qs,/ Objective

We formulate the problem as follows:

blah blah

## Thesis Structure

The thesis is structured to follow sections namely Literature Review, Dataset, Modelling, Experiments and Results, Conclusion, Appendix and Bibliography.

In Literature Review section will introduce the works that has already been carried out in this domain.

In the Dataset section a detailed insight to how our data is structured is elaborated.

In the Modelling section, we introduce the methodologies used to tackle the problem.

The Experiments and Results chapter gives an outlook on the outcomes of our work in addition to other findings we unearth.

Conclusion chapter summarizes the thesis briefly and would also give a small glimpse into areas of improvement.

Finally the Bibliography section lists out the articles cited for this thesis.

# Literature Review

There has been extensive research in modeling the Electric Motor with Convolution Neural Network (CNN) based on the images of the motor cross-section.

However our approach is progressive in the sense that once the KPIs are predicted we would like to be able to generate the inputs and reproducing images is not known to apt given the infamous known fact that AI generated images are faulty.

Instead by generating the parameters of the motor we can be rest assured of more precise results. Hence the need to focus on the inputs as they are with their parametric description.

Existing literature also covers works on modelling this work as tabular data using MLPs. Although this is fairly good forseeing the impact of generating the inverse process yet MLPs cannot necessarily learn all the intricacies within motor components.

Hence the need to better represent the data typically in the form of graphs and model Graph Neural Networks to achieve the desired results.

There has been close to no work of GNNs in this domain. However we see progress of GNNs in molecular chemistry and social networks usecases from which we draw inspiration.

# Dataset

Valeo an automotive company has supplied the dataset consisting of close to 1500 Double V Electric Motor parameters. There are close to 196 parameters which comprises of the geometric, physical and simulation properties of the motor.

The geometry of a whole Double V motor is as below

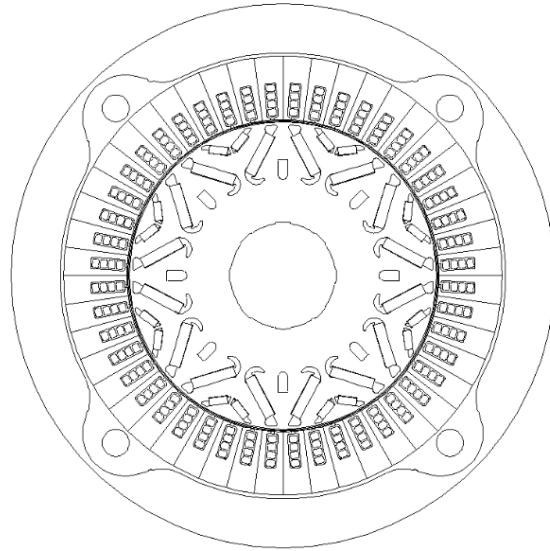


Figure 5: Complete EM Geometry(Source:Valeo)

Below is the geometry of 1/8 cross-section of the same motor.

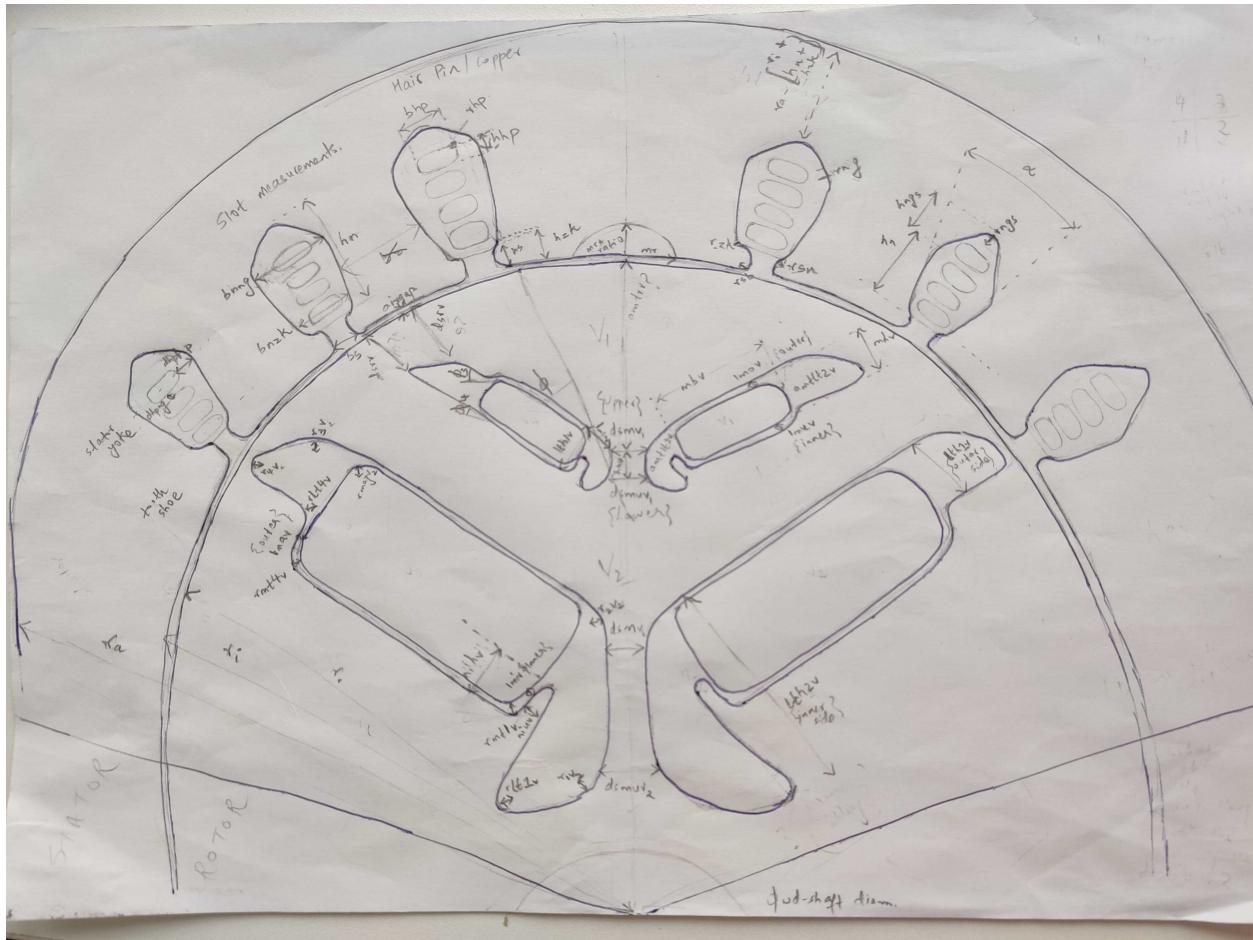


Figure 6: 1/8 Motor Crossection

Valeo has shared approximately 1500 .xlsx files for each motor variant. Each of the excel files contain multiple sheets. The sheets of interest to us are as below :

<b>Sheet</b>	<b>Sheet Name</b>	<b>Description</b>
Motor Parameters	input_data	<ul style="list-style-type: none"> <li>• Contains the input parameters for the motor model.</li> <li>• Includes geometric, physical, and simulation properties.</li> <li>• Unit dimensions if applicable are generally mm or degrees</li> </ul>
Speed Grid	NN	<ul style="list-style-type: none"> <li>• Contains the input parameters for the motor model.</li> <li>• Includes geometric, physical, and simulation properties.</li> <li>• Unit dimensions if applicable are generally mm or degrees</li> </ul>
Torque Grid	MM	<ul style="list-style-type: none"> <li>• Contains the torque grid.</li> <li>• Used for plotting the ETA grid.</li> </ul>
ETA Grid	ETA	<ul style="list-style-type: none"> <li>• Contains the ETA grid.</li> <li>• Have the same dimensions as that of NN and MM sheets</li> </ul>
Torque Curve	Mgrenz	<ul style="list-style-type: none"> <li>• Contains the values corresponding to torque curve.</li> <li>• Have the same columns corresponding to the speed grid.</li> </ul>

Table 1: Excel File Structure of an EM variant

## Data Preprocessing for MLP

For modelling the MLP, we present the data in tabular form with the parameters corresponding to columns. In order to make the data compatible with our model, some level of data processing was carried out as elaborated below.

### Deep Dive into the Input Parameters

All parameters including the additional ones in each topology is considered as a separate column and therefore if a particular column is topology dependent the data of the other topologies for that corresponding column is treated as 0 values.

The values are read and stored in their float equivalent to preserve data precision. Furthermore all degree columns are converted to their equivalent radian values to be fed to the model..EXPLAIN THE REASON WHY WITH CITATION PROBABLY:::

Parameter	Unit of Dimension	Mean	Standard Deviation	Value Range	Single V Magnet Topology	Double V Magnet Topology	Nabla Magnet Topology
<b>General Parameters</b>							
N	#	99	50–150	100	110	90	×
simQ	#	99	50–150	100	110	90	✓
r_a	mm	99	50–150	100	110	90	1,400
r_i	mm	99	50–150	100	110	90	1,400
<b>Rotor Parameters</b>							
Milan	M	1,500	1,000–2,000	1,500	1,600	1,400	1,400
Julie	F	1,325	1,000–1,500	1,300	1,350	1,325	1,400
<b>Stator Parameters</b>							
Sekhar	M	922	800–1,100	900	950	915	1,400
Dipen	M	598	500–700	600	625	570	1,400

Table 2: EM Input Parameters

### Deep Dive into the 2 Dimension (2D) KPI(Torque Curve)

Below is the standard deviation of the test dataset for the 2D KPI(Torque Curve).

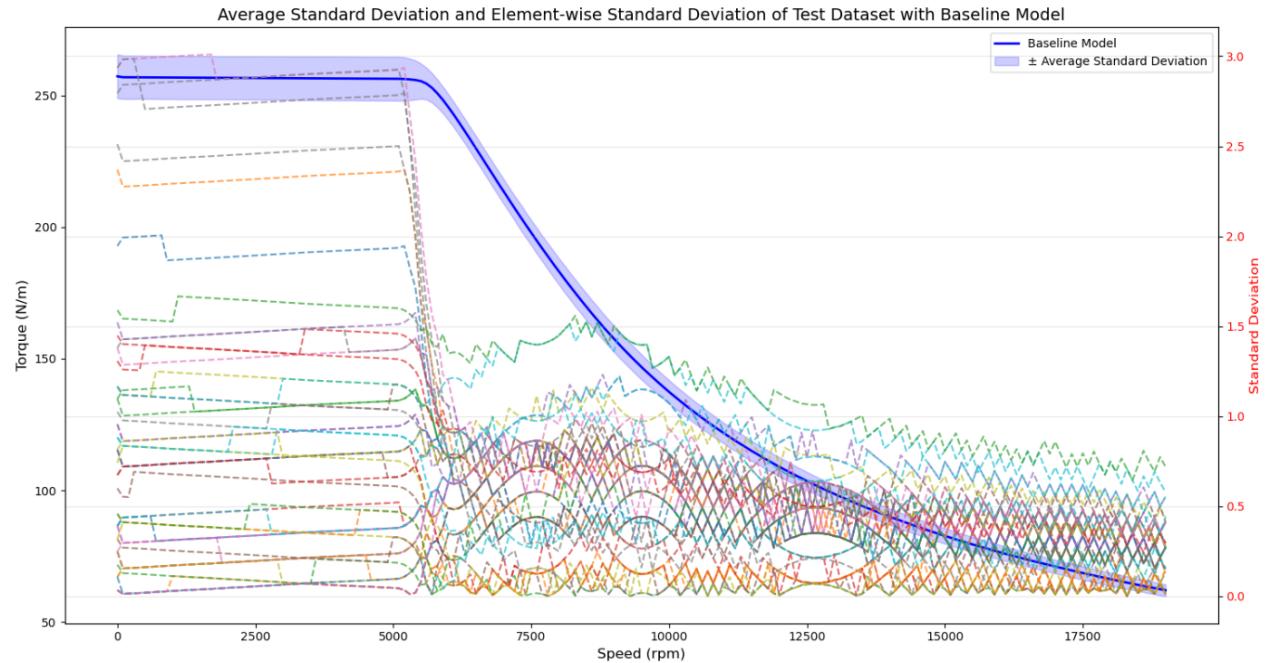


Figure 7: Standard Deviation of 2D KPI(ETA) Targets

### Deep Dive into the 3 Dimension (3D) KPI(ETA Curve)

From the above plot, we can conclude that the curve is relatively constant for 1/4 of the curve after which it is monotonically decreasing. This finding is critical for how we modelled the loss regularization for the

target. It will be further elaborated in the Section CITE LOSS FUNCTION

As the target values Mgrenz and ETA grids are not provided with the correct dimensions we have an additional step which takes the maximum torque value from the Mgrenz grid and create a similar grid ranging from -maximum torque to maximum torque.

We then choose only the rows corresponding to this range from the actual MM grid supplied and the same row indices is used to retrieve the ETA grid.

This step ensures that we grant the model the correct dimensions of the ETA grid based on Torque curve and predict likewise.

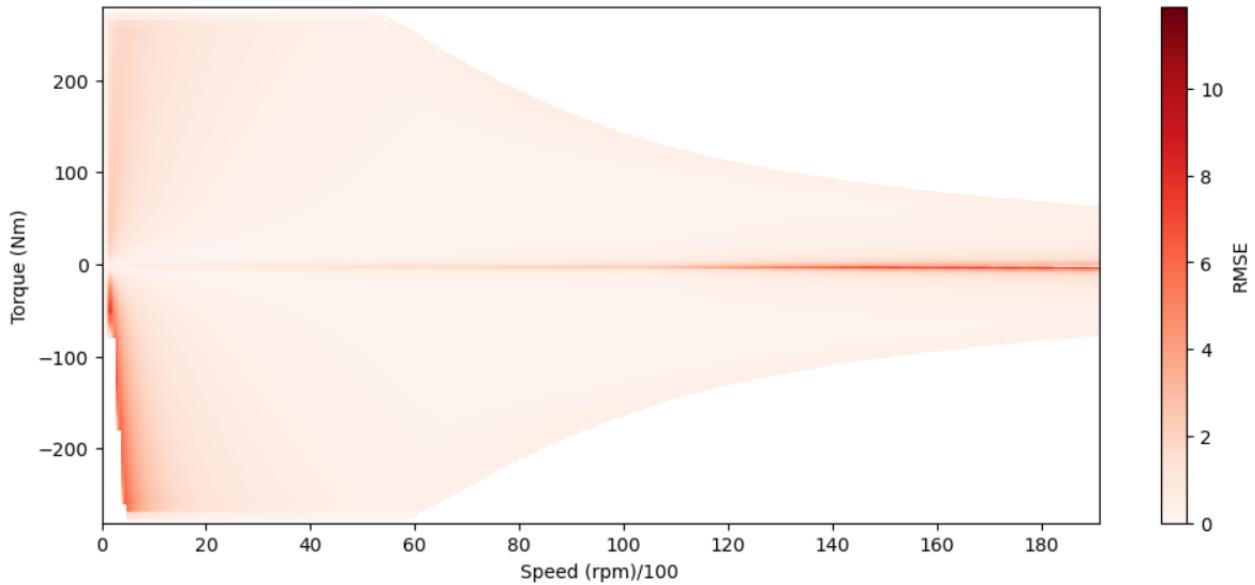


Figure 8: Standard Deviation of 3D KPI(ETA)

More apparent below :

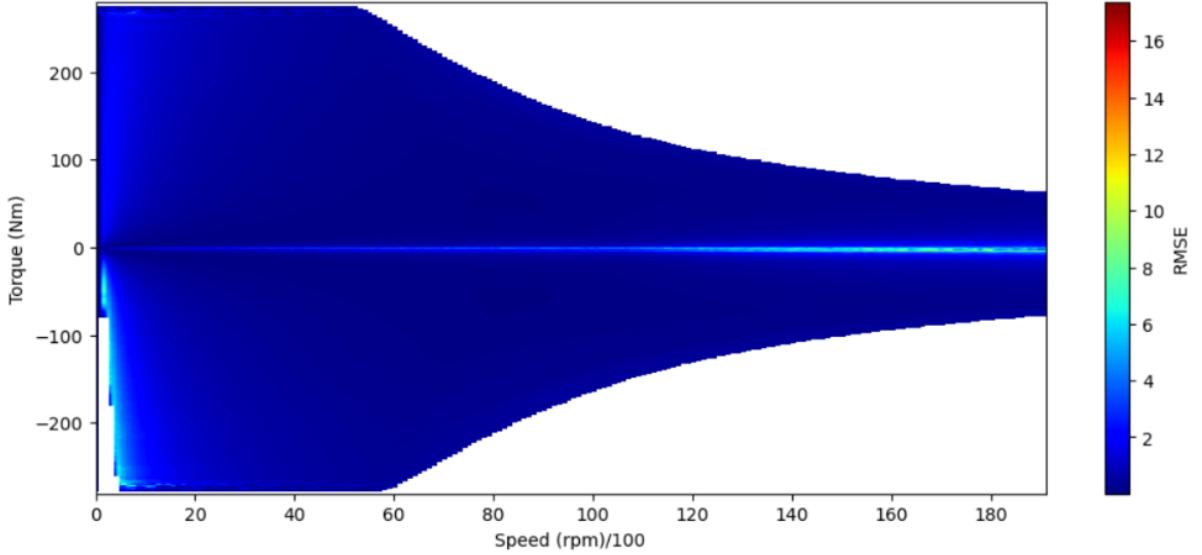


Figure 9: Standard Deviation of 3D KPI(ETA)

From our analysis of the above plots on the standard deviation of the ETA grids from the test dataset, we notice that the data shared was faulty in the ETA corresponding to negative torque values. As these are FEM simulations, it is probably an effect of a post processing step taken by the Motor builder. The ETA grid corresponding to negative torque values when motor is in generating mode is not the replica as how it is for positive torque values when motor is in monitoring mode. This is evident from low speed high torque distribution area where we can see NAN values across the test dataset

Additionally the ETA grid envelope is completely dependent on its equivalent torque curve. The area beneath the boundary of which is looked into by the EM manufacturers to determine the car's efficiency in the operating cycle.

From the plots received we find the ETA grid does not only cover the area within the torque curve but also a slightly exaggerated area typically in the portion where the 1/4 where the curve starts to monotonically decrease. To ensure that our model is not tricked into assuming predictions for this extra area we extract the grid to be of the same shape as that of its curve. This is a step also done as part of data post processing to be further covered in Section BLAH CITE.

Based on this finding, we have decided to only consider the ETA grid grid in monitoring mode for predictions and mirror the same to replicate the efficiency when it is in generating mode.

As reading the files take up a lot of time and compute, we read the files as a one-time job at the start and store them into pythonic objects for faster access in the future.

Both the input and target values for 1st KPI is stored locally as pandas dataframe .csv files whereas the 2nd KPI is stored as separate csv files per variant considering it is in the form of a 2D array.

The csv files are then concatenated and stored into a numpy array conserving dimensionality by padding nan values to match dimensionality of the largest 2D array among them.

The array is then saved locally as a .npy file for easy access and loading during training.

The input data consists of about 1500 examples mostly from the Double V Magnet Topology (and about 3 examples each for the other 2 topologies).

## Scaling

We have used Standard Scaling for the input and both the outputs.

The Scaling is formulated as shown below :

$$z = \frac{x - \mu}{\sigma}$$

where

- $x$  : Input
- $\mu$  : Mean
- $\sigma$  : Standard Deviation

both Mean and Standard deviation are calculated across columns. This is attributed to the fact we have columns with different ranges for the input. Meanwhile the torque curve and the ETA grid are flattened to be 1D and thus have only 1 column.

We tried scaling with only the inputs but this resulted in loss not converging as we thought. It would probably take more runs for the loss to converge and could be taken care of in future as improvements.

## Dataset splitting

We have also split the dataset to have about 50 samples for test and the remaining is used for 5 fold cross validation with 80:20 split for training and validation.

Within 5 folds, we expect to cover most grounds on training and have good monitoring on the model's performance for each fold. The best performing model is then chosen from the folds based on the model which has a combination of least aggregated loss for both outputs and scores for each output which are closest to 0.

# Graph Modelling

## Data Preprocessing for GNN

### Node types

#### 1. General

- General parameters:

$$r = \{r_i\} \quad \forall i \in \{a, r, o\}$$

where:

- $r_a$ : Outer Radius of the Stator
- $r_r$ : Outer Radius of the Rotor
- $r_o$ : Center of the EM

#### 2. Stator

- Slot windings:

$$sw = \{s_i w_j\} \quad \forall i \in \{1, \dots, QSim\}, \quad \forall j \in \{1, \dots, N\}$$

- Slots:

$$s = \{s_i\} \quad \forall i \in \{1, \dots, QSim\}$$

where

- $Qsim$ : Count of slots in the Stator
- $N$ : Count of copper windings per slot

#### 3. Rotor

- Magnet Flux Barriers:

$$v = \{v_{ij}\} \quad \forall i \in \{1, \dots, T\}, \quad \forall j \in \{1, \dots, V\}$$

- Magnets:

$$vm = \{v_i m_j\} \quad \forall i \in \{1, \dots, T\}, \quad \forall j \in \{1, \dots, V\}$$

where

- $T$ : Topology type of the EM
- $V$ : Type of Magnet

As Valeo only manufactures Double V magnets we consider it to be 2

### Edge types

#### 1. Angle

##### Relevant Paths

$$vm - - vm = \{v_{i_1} m_{j_1} - v_{i_2} m_{j_2}\} \quad \forall i_1, i_2 \in \{1, \dots, T\}, \quad \forall j_1, j_2 \in \{1, \dots, V\} \mid i_1 = i_2, \quad j_1 \neq j_2$$

`angle=vm-vm`

## 2. Distance

### Relevant Paths

$$\begin{aligned}
vi--vi &= \{v_{ij_1} - v_{ij_2}\}, \forall i \in \{1, \dots, T\}, \forall j_1, j_2 \in \{1, \dots, V\} \mid j_1 \neq j_2 \\
vi--vj &= \{v_{i_1j} - v_{i_2j}\}, \forall i_1, i_2 \in \{1, \dots, T\}, \forall j \in \{1, \dots, V\} \mid i_1 \neq i_2 \\
v--vm &= \{v_{ij} - v_{im_j}\} \forall i \in \{1, \dots, T\}, \quad \forall j \in \{1, \dots, V\} \\
v--rr &= \{v_{ij} - r_r\}, \forall i, j \in \{1, \dots, T\} \\
o--r &= \{(o - r_r), (o - r_a)\} \\
rr--s &= \{r_r - s_i\}, \forall i \in \{1, \dots, QSim\} \\
s--sw &= \{s_i - s_i w_j\}, \forall i \in \{1, \dots, QSim\}, \forall j \in \{1, \dots, N\} \\
s--ra &= \{s_i - r_a\}, \forall i \in \{1, \dots, QSim\} \\
sw--sw &= \{s_i w_{j_1} - s_i w_{j_2}\}, \forall i \in \{1, \dots, QSim\}, \forall j \in \{1, \dots, N\} \mid (j_1 == j_2 - 1) \\
\text{distance} &= vi-vi + vi-vj + v-vm + v-rr + o-r + rr-s + s-sw + s-ra + sw-sw
\end{aligned}$$

### Node Features

1. **v** = {lmssov, lth1v, lth2v, r1v, r11v, r2v, r3v, r4v, rmt1v, rmt4v, rlt1v, rlt4v, hav}
2. **vm** = {mbv, mhv, rmagv}
3. **r** = {r}
4. **s** = {b\_nng, b\_nzk, b\_s, h\_n, h\_s, r\_sn, r\_zk, r\_ng, h\_zk}
5. **sw** = {bhp, hhp, rhp}

### Path Features

1. **vm–vm** = {deg\_phi}
2. **vi–vi** = {dsm, dsmu}
3. **vi–vj** = {amtrvj-amtrvi}
4. **v–vm** = {lmav, lmiv, lmov, lmuv}
5. **v–r** = {amtrv, dsrv}
6. **o–r** = {r}
7. **rr–s** = {airgap}
8. **s–sw** = {dhphp}
9. **sw–sw** = {dhpng}
10. **s–ra** = {r\_a-(r\_i + h\_n + h\_zk)}

The heterogeneous graph that was constructed earlier is as below:

## Heterogeneous GNN Model

We find the heterogeneous graph to be most apt for our use case with its different node and edge types as it preserves both the structural and semantics of our data.

This property is crucial in modelling our use case as we will then have similar node-edge types per topology. In addition the count of certain parameters with the motor such as stator poles with its corresponding slot and rotor magnets is made more comprehensible to the model having new nodes and edges whereas for the MLP architecture this information is represented only as a number in yet another column.

Heterogeneous GNN generally work by having separate non linear functions convolve over each edge type during message computation and over each node type when aggregating the learned information.

Visualisation of an Electric Motor as a Heterogeneous Graph

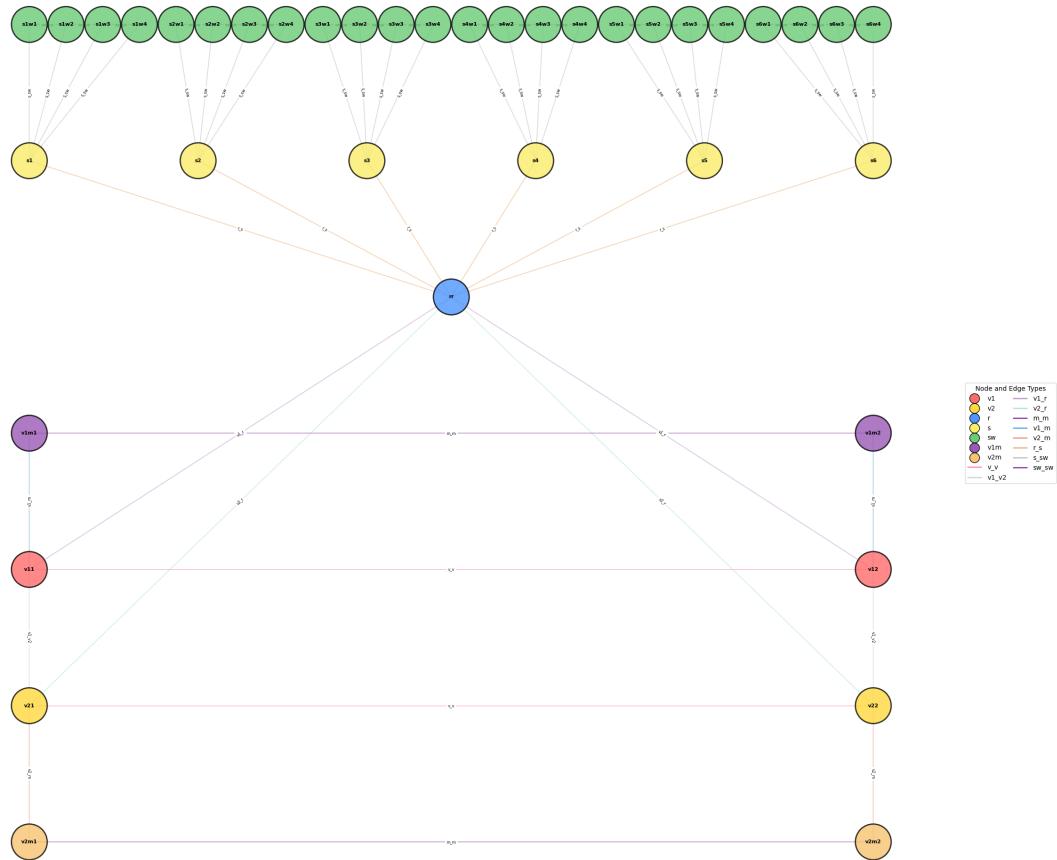


Figure 10: HetGraph

# Modelling & Evaluation

Since we aim to predict continuous vector values, we model this task into a regression problem. As a baseline, we first train a MLP on the tabular representation of the data and work on it further to do the same with a heterogeneous GNN.

## MLP Model

For the MLP model, we use a single model with input features corresponding to all the features in the tabular topology invariant representation of the data.

The model architecture is build to predict both the 2D and 3D KPIs by having 2 separate output layers for each of the KPIs.

Since the 2D KPI's targets are relatively learnable than that of the 3D KPI's targets we have experimented with fewer feed forward layers in the former than in the latter. RELU layers were also added in between to serve as the activation function and produce non-linearities in the model.

Dropout layers ensure that not all neurons in each layer are used up during training to prevent the model from memorizing the data and hence overfitting. Batch normalisation layers are used to normalize the input to the next layer and hence speed up the training process.

The model architecture is designed to have 2 fully connected layers with RELU as an activation function and a 1D batch normalisation layer with a dropout layer for the 1st output.

We have also used a dropout percentage of 0.2 to ensure the model does not overfit the data by dropping out 20 percent of the neurons in the layer. Although the targets for the first output are an array of integer values, we use the float tensor and not integer tensor to represent the data else it would become a classification problem and not a regression problem as it should be.

For the second output we have 5 fully connected layers with RELU as an activation function and dropout and 2D batch normalisation for each layer.

## Loss Functions

The Mean Squared Error (MSE) loss is the loss function used for our problem with the intention that the losses are minimized. We have adopted 2 methodologies to regularize the loss one each for the 2D and 3D KPI.

### Loss Regularization for 2D KPI(Torque curve)

The MSE loss for the 2D KPI is formulated as shown below :

$$\text{Y1 Loss} = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} \sum_{j=1}^h (y_{ij} - \hat{y}_{ij})^2$$

where

- $n$  : EM samples
- $h$  : dimensionality of 1D vector

To smoothen out the curve for the 2D KPI(Torque curve) we apply a loss regularisation factor to take into account before backpropagating it the model during training.

We observed the curve inherently follows a decreasing pattern CITE AGAIN::COULD BE A KNOWN FACT:::and hence using this knowledge penalize the loss for non-decreasing values within each prediction. CITE!!!!If the electromagnetic coil is enabled by the commutator for the time span t3, the (almost) maximal current is running through it's loops and the (almost) maximal magnetic field strength is generated. The (almost) maximal torque is acting on the rotor. If the time span is shortened to t2 by increasing rotational speed, a slightly lower torque is acting, because the current through the coil is decreasing slightly. When reducing the time span to t1, the coil gets disconnected from the input voltage even though just half the maximum current is reached. Accordingly the torque decreases significantly:

$$Y1 \text{ Loss Regularization} = \frac{1}{n} \sum_{i=1}^n \frac{1}{h-1} \sum_{j=1}^{h-1} (\text{ReLU}(\hat{y}_{ij+1} - \hat{y}_{ij}))^2$$

where

- *RELU : Activation Function*

Theoretically, we would expect the model to generate better predictions but on closer observation we notice the curve is still not smooth.

A reason to attribute this could be the model's incapability to infer that loss decrease depends on not just the prediction and target values but also within prediction values. Our deduction is that a single value calculated for the entire curve may not be sufficient to regularize this loss as we imagined.

A reason to attribute this could be the model's struggle to attempt to generate the next point in the curve to be smaller than the previous point yet also for the further point in addition to making sure the MSE is calculated accurately. Our hypothesis is this loss regularization creates a tug of war between the MSE loss and the regularization term when predicting the whole curve and hence the ever fluctuating curve being generated.

### **Loss Customization for 3D KPI(Efficiency Grid)**

The MSE loss is calculated for the 3D KPI is formulated as shown below :

$$Y2 \text{ Loss} = \frac{1}{n} \sum_{i=1}^n \frac{1}{w} \frac{1}{h} \sum_{j=1}^w \sum_{k=1}^h (M_{ijk} \cdot y_{ijk}) - (M_{ijk} \cdot \hat{y}_{ijk})^2$$

where

- $M_{ijk}$  : Mask matrix

$$M_{ijk} = \begin{cases} 1 & \text{if } y_{ijk} \neq -1 \\ 0 & \text{if } y_{ijk} = -1 \end{cases}$$

- $w$  : 1st dimension of 3D vector
- $h$  : 2nd dimension of 3D vector

The ETA Grid is a 3D plot of real numbers ranging between 0 and 100.NEED TO CITE...EVEN IF IT IS A KNOWN FACT We noticed in some portions of the grid, the plot not visible as it had nan values.

As ANN cannot be trained to predict NAN values we had these replaced to -1 and have a binary mask constructed such that values corresponding to -1 in the target have value 0 and all other values as 1.

The mask is then multiplied with both the target and the prediction. After this step, the MSE loss is calculated and backpropagated.

Mathematically, this process can be expressed as above

This formulation ensures that the -1 values (which replaced NaN values in the grid) are ignored in the loss calculation, as they are multiplied by 0 in the mask.

It is a known fact that at 0 Torque, the corresponding efficiency values for the motor is 0. To model this conception into our network, we have attempted to retrieve the center row of our ETA grid and the values it consists finally determine the violation count. We formulate is mathematically as below:

$$r_{\text{mid}} = \left\lfloor \frac{w}{2} \right\rfloor \quad \forall i \in \{0, \dots, n - 1\}$$

$$\text{Mid} = Y_{:, r_{\text{mid}} : j}$$

$$\text{Y2 Loss Regularization} = \frac{1}{h} \sum_{j=1}^h |\text{Mid}_j|$$

We take the absolute values from the center row of the matrix else the loss would be negative and could negate with positive values. This would have a detrimental effect when calculating the loss

Total Loss = Y1 Weight  $\times$  (Y1 Loss + ( $\lambda_{y1} \times$  Y1 Loss Regularization)) + Y2 Weight  $\times$  (Y2 Loss + ( $\lambda_{y2} \times$  Y2 Loss Regularization))

where

$$\lambda_{y1} : \text{Y1 Loss Regularization Weight}$$

$$\lambda_{y2} : \text{Y2 Loss Regularization Weight}$$

- Y1 Weight : Y1 Loss Weightage
- Y2 Weight : Y2 Loss Weightage

## Optimizer

Adam optimizer is used for optimization as it is known to be computationally efficient and requires little memoryBLAH BLAH

The optimizer acts once the loss is backpropagated across training each batch of the dataset.

We have also experimented with a learning rate scheduler which reduced the learning rate exponentially by a gamma parameter to decay learning as training progresses across epochs.

## Evaluation Metrics

The evaluation metrics we have considered for our regression problem is the root mean squared error. The model with the least combined loss and prediction scores closest to 0 is ideal for our application.

### Evaluation Metrics for 2D KPI

The y1 loss for the 2D KPI is formulated as shown below :

$$\text{Y1 score} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{1}{h} \sum_{j=1}^h (y_{ij} - \hat{y}_{ij})^2}$$

where

- n : EM samples
- h : dimensionality of 1D vector

## Evaluation Metrics for 3D KPI

The  $y_2$  loss for the 3D KPI is formulated as shown below :

$$Y2 \text{ score} = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{w} \frac{1}{h} \sum_{j=1}^w \sum_{k=1}^h (y_{ijk} - \hat{y}_{ijk})^2}$$

where

- $n$  : EM samples
- $w$  : 1st dimension of 3D vector
- $h$  : 2nd dimension of 3D vector

## Post Processing

The mean and standard deviation from the train-validation datasets are applied to transform the test dataset to maintain uniformity on the predictions generated. In the case of new files we first convert it into the tabular representation our model consumes and then do the scaling.

This is the reason why we preserve the same scalers used during training as we not only evaluate our dedicated test dataset but also for clients to use on demand.

The inverse scaling is formulated as shown below :

$$x = \frac{z \times \sigma}{\mu}$$

where

- $z$  : Scaled Input
- $\mu$  : Mean
- $\sigma$  : Standard Deviation

# Experiments and Results

## Experiments with MLP

The hyperparameters were chosen via a random grid search and we finalized those that resulted in the ideal scores for both  $y_1$  ad  $y_2$  respectively. The hyperparamters for the model was tuned over observations of the model's performance across 5 fold cross validation training.

The splits are saved locally and can be used to later to ensure reproducibility.

Hyperparameters	Description	Value
lr	Learning Rate	1.927
n	Count of Hidden Layers	0.4962
d	Dimensionality of Hidden Layers	1.927
gamma	Learning Rate Scheduler	0.4962
n	Batch Size	1.927
n	Number of Epochs	0.4962
p	Dropout Probability	1.927
lambday1	Y1 Loss Regularizer	0.4962
lambday1	Y1 Loss Regularizer	0.4962
w1	Weightage of Y1 Loss	0.4962
w2	Weightage of Y2 Loss	0.4962

Table 3: Hyperparameter Tuning

Over the 5 folds, the model performance is observed to ensure its stability and the best performing model is saved to be loaded later for performance. The criteria for choosing the best performing model is roughly estimated based on the least combined loss and the combined score closest to 0.

We chose Weights & Biases (Wandb) to log metrics from the training run and to monitor model performance across folds.

Below is the visualisation of the training and validation metrics for both KPIs.

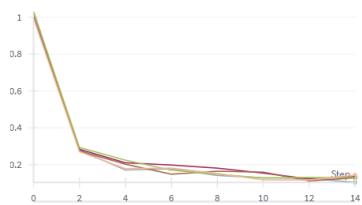


Figure 11: Aggregated Training Loss

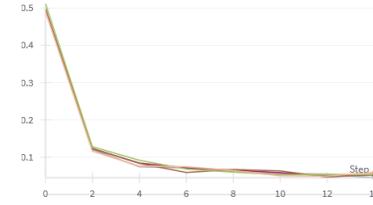


Figure 12: Training Loss for Torque Curve

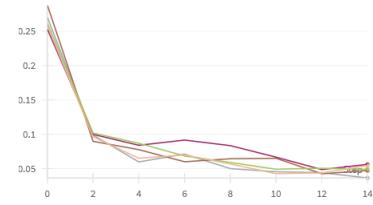


Figure 13: Training Loss for ETA grid

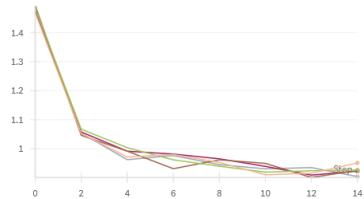


Figure 14: Aggregated Training Score

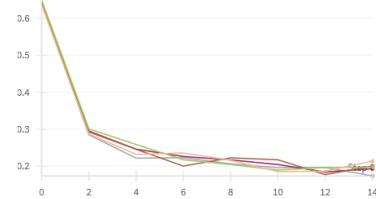


Figure 15: Training Score for Torque Curve

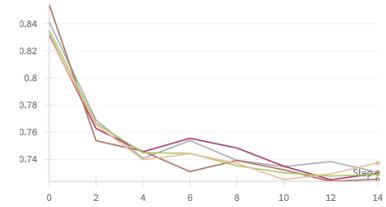


Figure 16: Training Score for ETA grid

From the training plots we see that the model has converged after having run for 10 epochs with a learning rate of 0.0075.

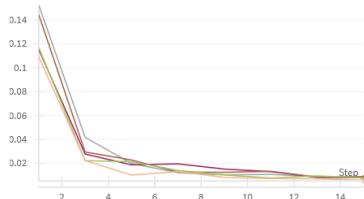


Figure 17: Aggregated Validation Loss

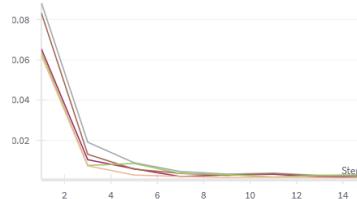


Figure 18: Validation Loss for Torque Curve

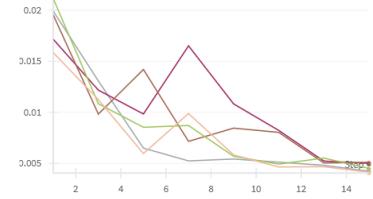


Figure 19: Validation Loss for ETA grid

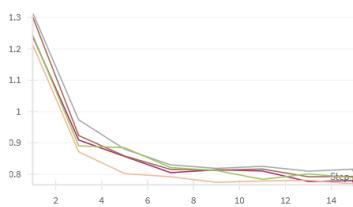


Figure 20: Aggregated Validation Score

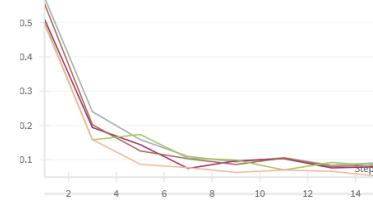


Figure 21: Validation Score for Torque Curve

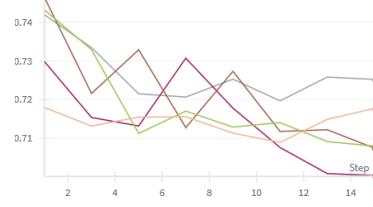


Figure 22: Validation Score for ETA grid

We see a good fit of the model to the data with corresponding y1 and y2 scores approaching close to 1. We have also enabled saving the trained model locally so it can be loaded on demand by the client to run inference.

We have narrowed down scoring to follow the below criteria:

Scoring	Excellent	Very Good	Good	Average	Bad
Y1	0-1	1-2	2-4	4-6	>6
Y2	0-0.5	0.5-1	1-2	2-4	>4

Table 4: Scoring Criteria

## Results with MLP

The results of the MLP model from inference is as below:

### 2D Torque Curve Results with MLP

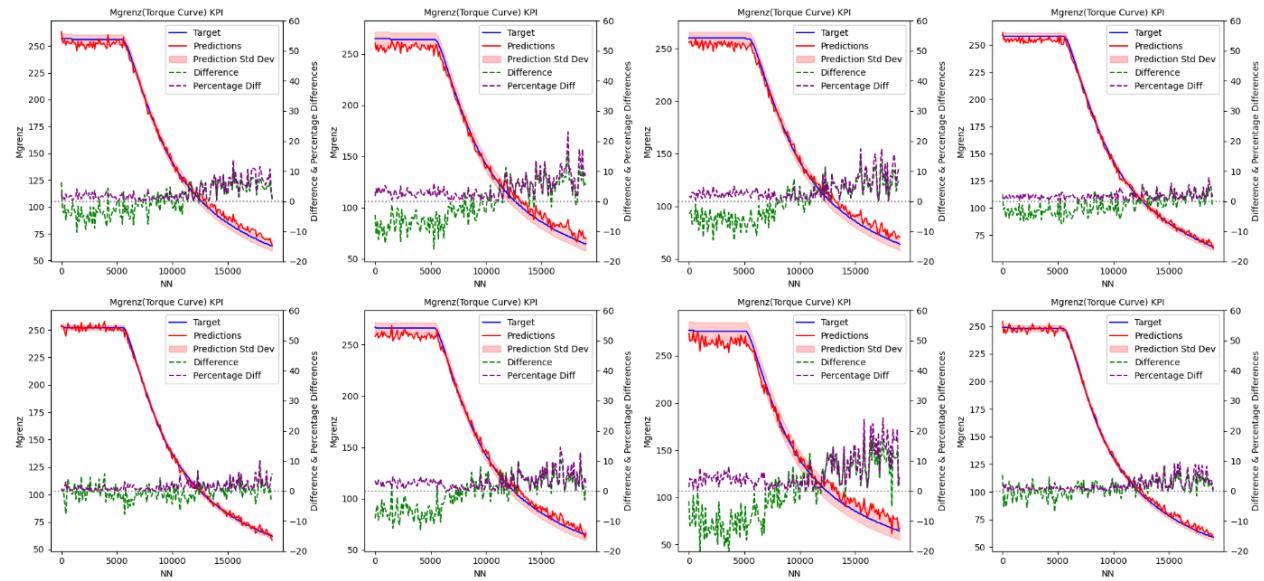


Figure 23: MLP Training Results for 2D KPI(Mgrenz)

The Average RMSE and element wise RMSE for the test dataset performance with the MLP is as below:

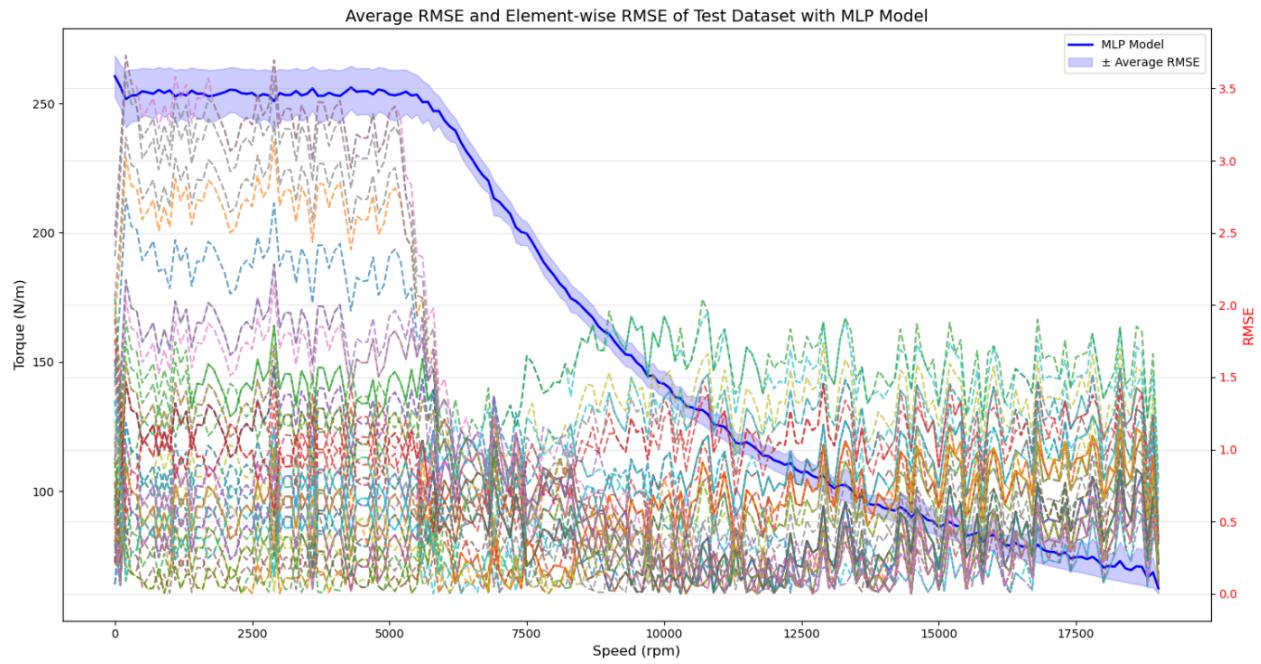


Figure 24: Test Dataset RMSE Evaluation for 2D KPI(Mgrenz)

We also have the score statistics of how the model performs over the test dataset.

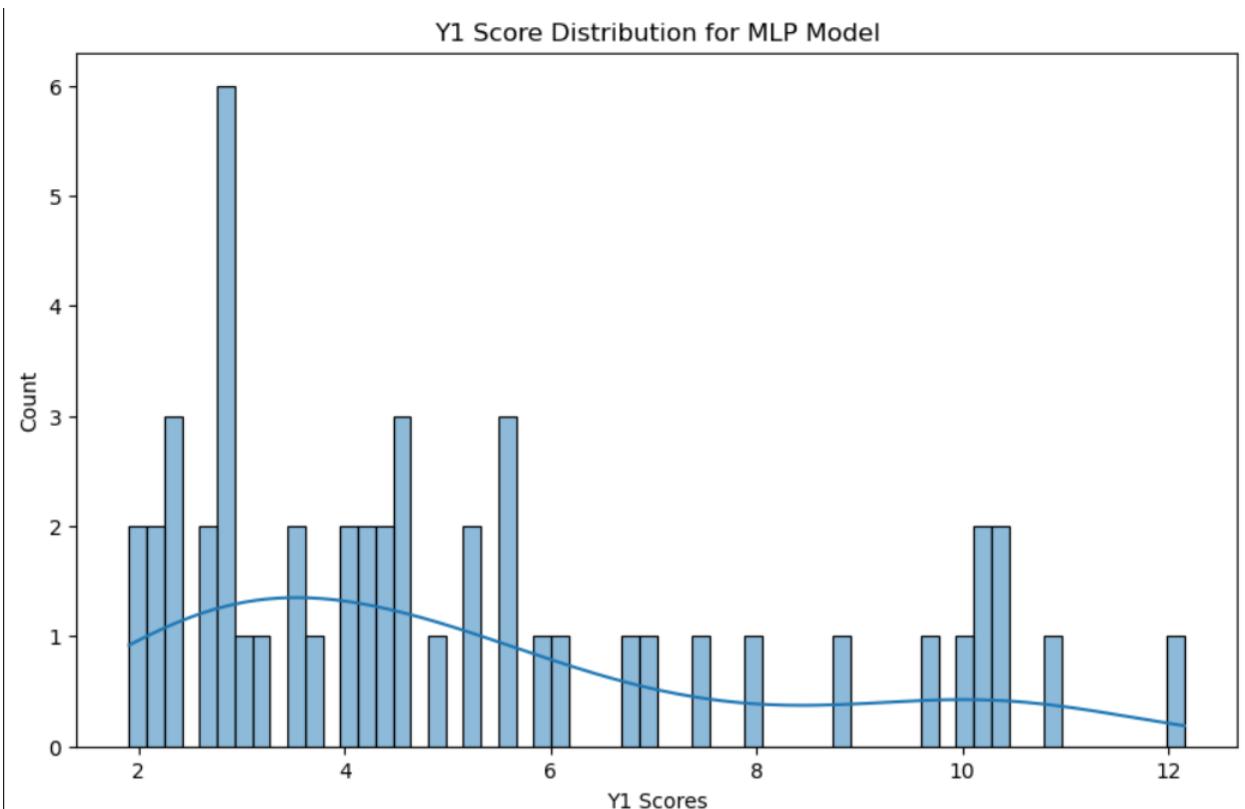


Figure 25: MLP Score statistics for 2D KPI(Mgrenz)

### 3D ETA Grid Results with MLP

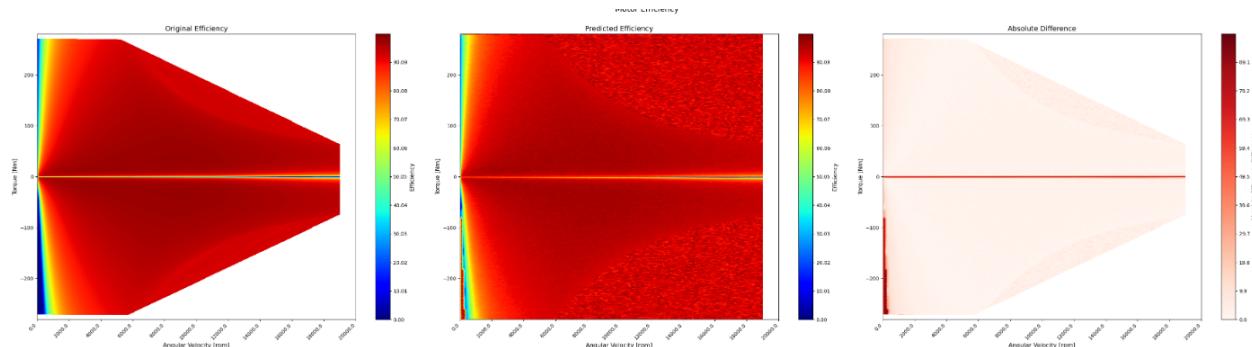


Figure 26: 1st MLP Training Results for 3D KPI(ETA)

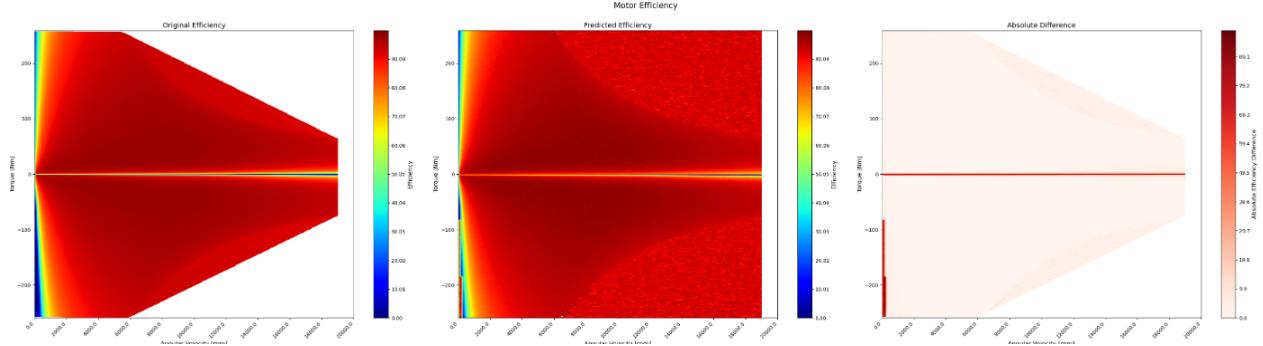


Figure 27: 2nd MLP Training Results for 3D KPI(ETA)

Since we are predicting a padded matrix to ensure dimensionality sync different ETA grids, the grid contains values even outside the boundary of the ETA envelope. Hence, we attempted to slice the shape of the Torque curve from the ETA grid by counting the number of columns a row to have based on consecutive values in the curve. As you can see below, this has resulted in clipping some of the actual ETA envelope with the EM's operating drive cycle.

This brings us back to the point that it is imperative the prediction of the Torque curve is as good as possible to the actual curve as the envelope of the ETA grid is inherently dependent on it.

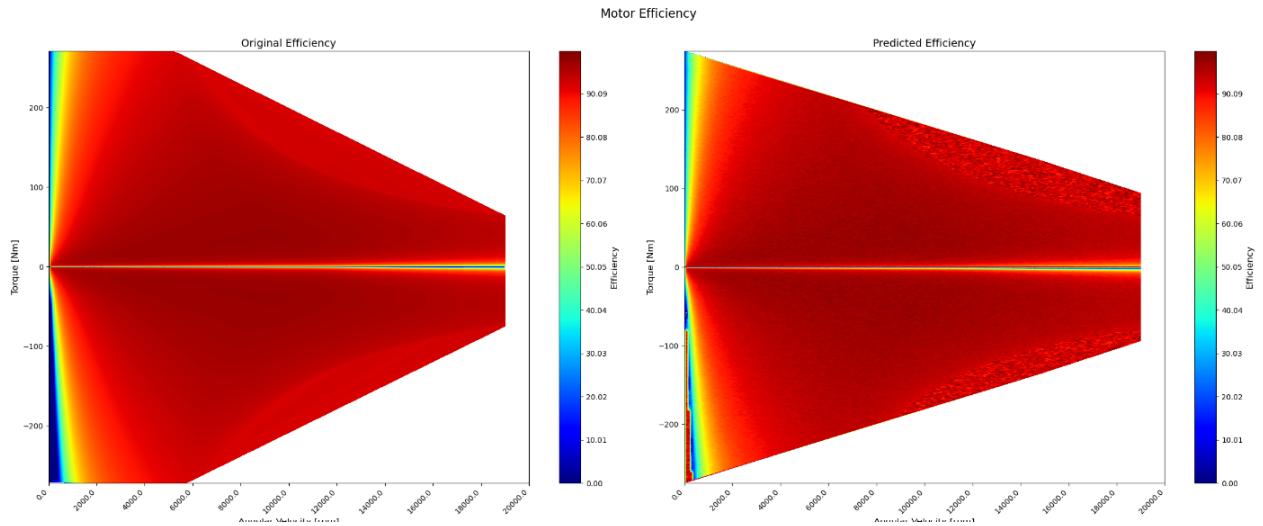


Figure 28: 3rd MLP Training Results for 3D KPI(ETA)

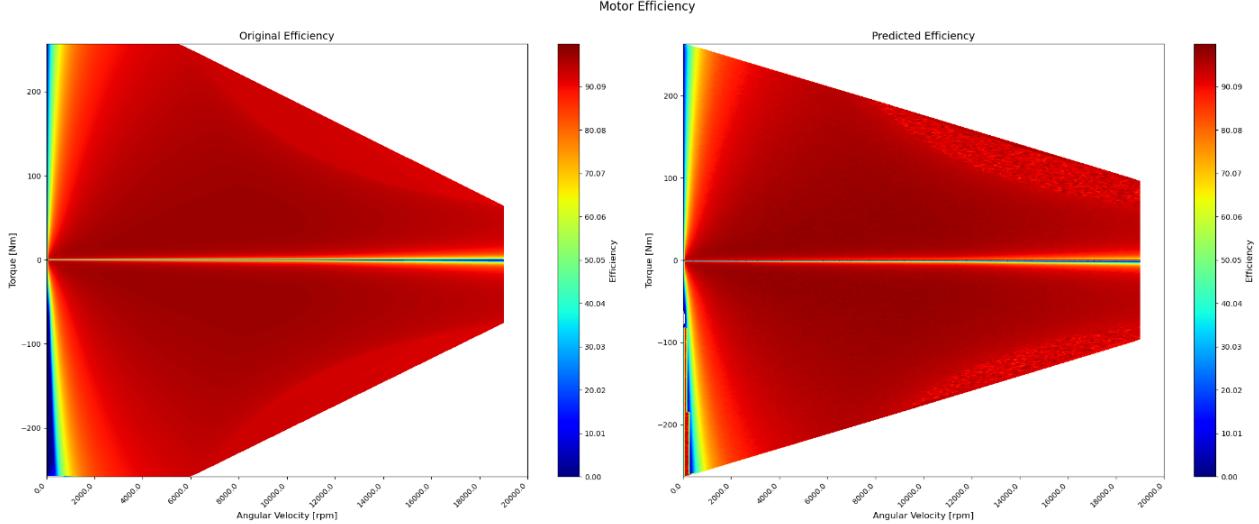


Figure 29: 4th MLP Training Results for 3D KPI(ETA)

Observations from the predictions helped to correct few discrepancies in our development for instance in the ETA grid we replaced 0 with NAN values which we later understood were both represented different in the grid.

As efficiency values can take up values only between 0 and 100, we consider the same as constant across plots and use it as a baseline for determining the levels in the contour plot.

We have also left the output predictions for the Torque curve to remain as float values even when the target values are integers to preserve data precision. We give the client the flexibility to turn this on/off demand. We also have the score statistics of how the model performs over the test dataset.

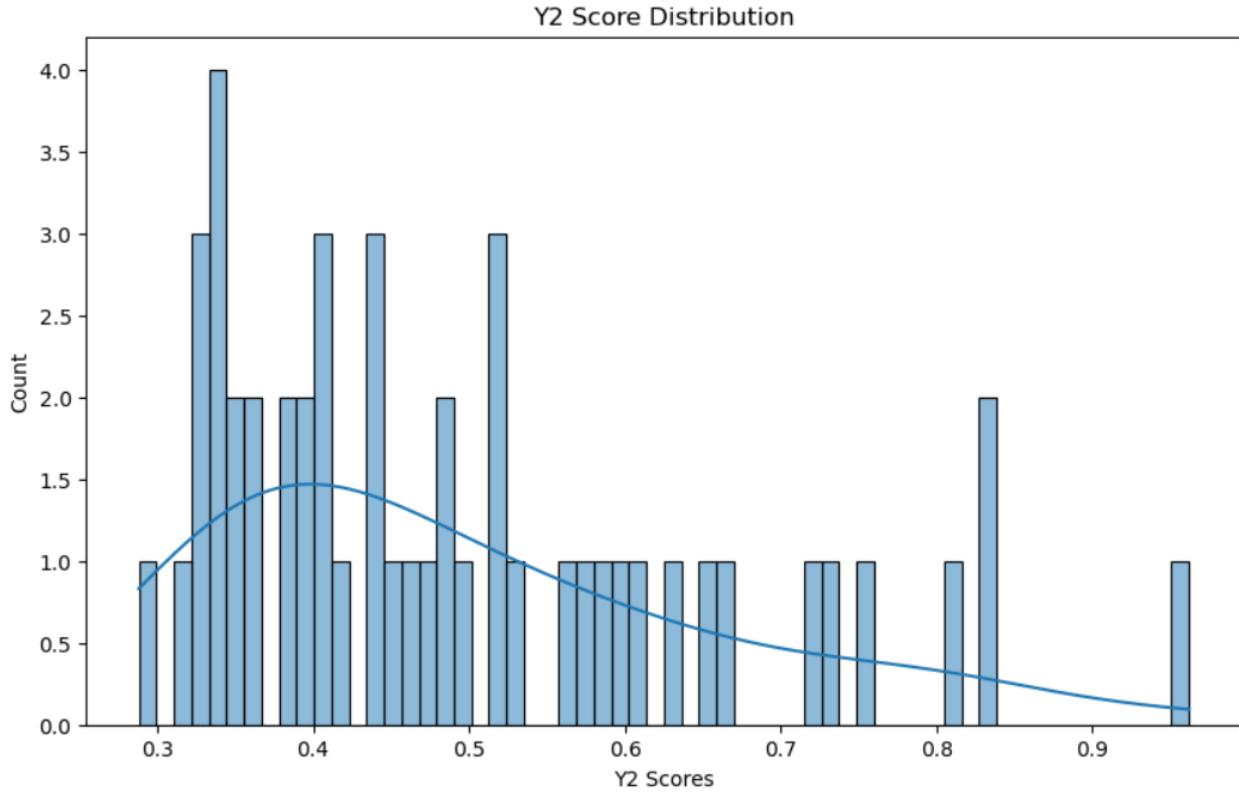


Figure 30: MLP Score statistics for 3D KPI(ETA)

We have used python 3.12.2 for our development and the pytorch library compatible with cuda. The model was trained on a NVIDIA V100 GPU with blah blah.

## Results with Baseline

From our observations of how the predictions closely resembled that of the target values, we have proceeded with a baseline model which is essentially the average of the target in all ground truth. Based on the standard deviation of the target values with the baseline as is formulated in Equation CITE, the corresponding plots are generated.

$$\text{Y1 Element wise RMSE} = \sqrt{\frac{1}{n}(\bar{x} - x_i)^2} \quad \forall i \in \{0, \dots, w-1\}$$

where

- $n$  : EM samples
- $w$  : 1st dimension of 3D vector
- $\bar{x}$  : Baseline Average mean

The results of the Baseline model from inference is as below:

## 2D Torque Curve Results with Baseline

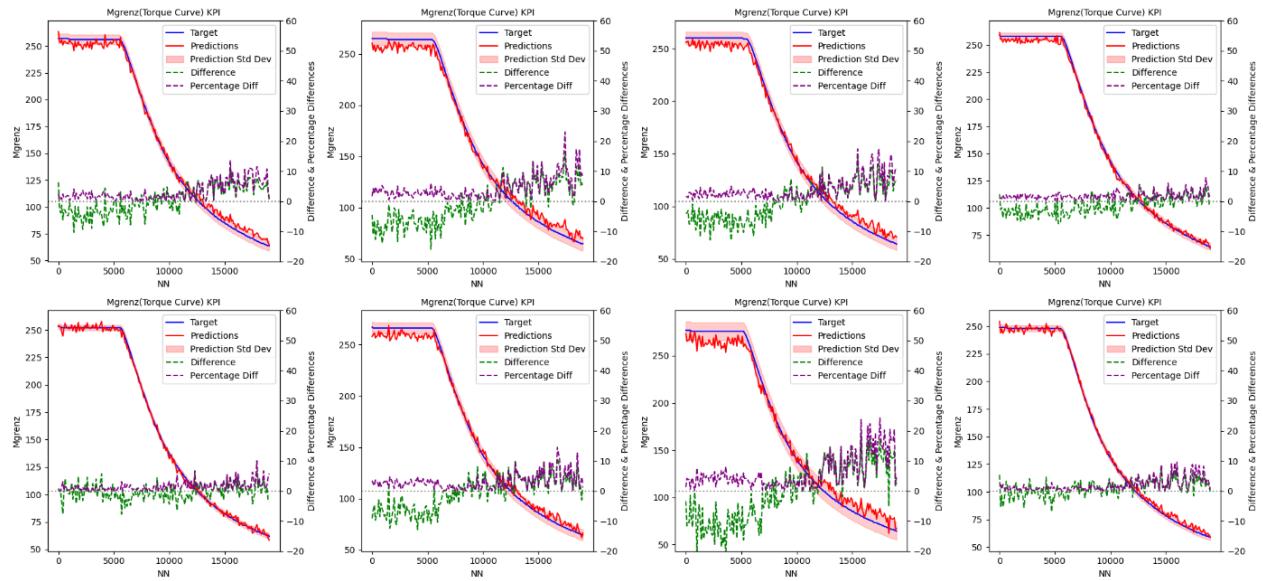


Figure 31: Baseline Results for 2D KPI(Mgrenz)

We also have the score statistics of how the Baseline model performs over the test dataset.

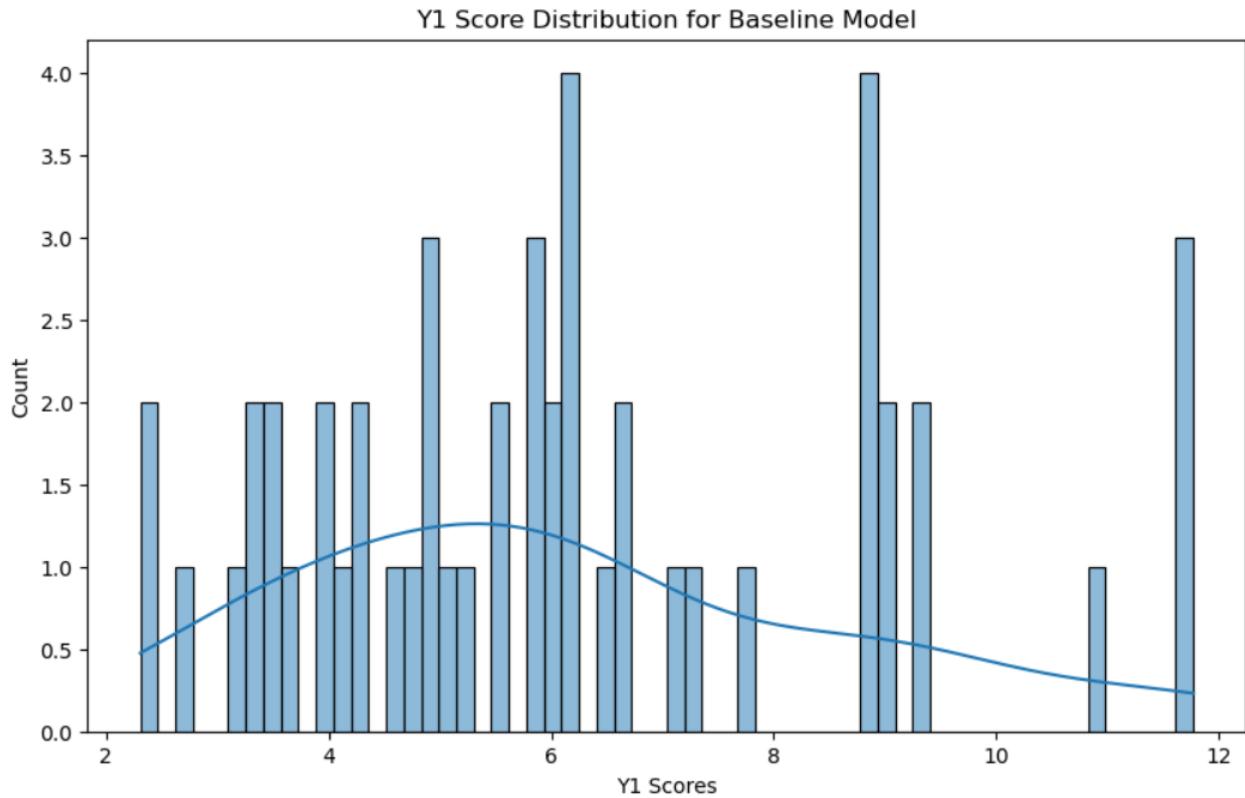


Figure 32: Baseline Score statistics for 2D KPI(Mgrenz)

### 3D ETA Grid Results with Baseline

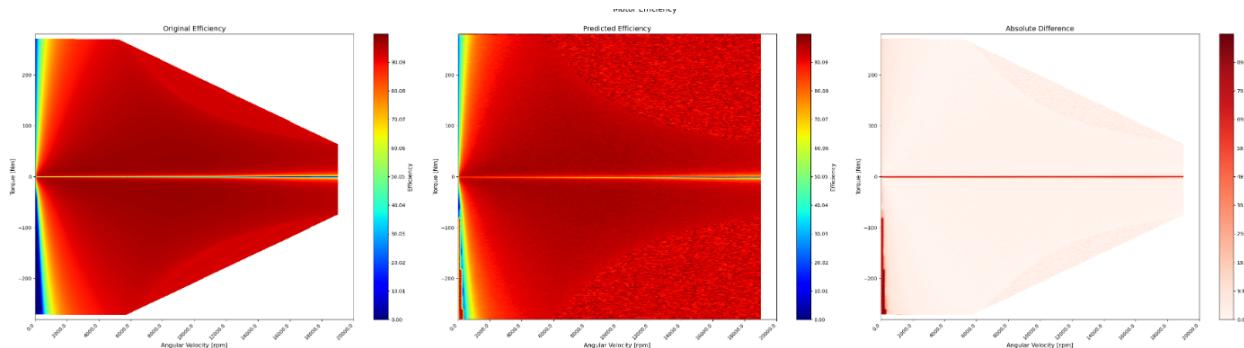


Figure 33: 1st Baseline Results for 3D KPI(ETA)

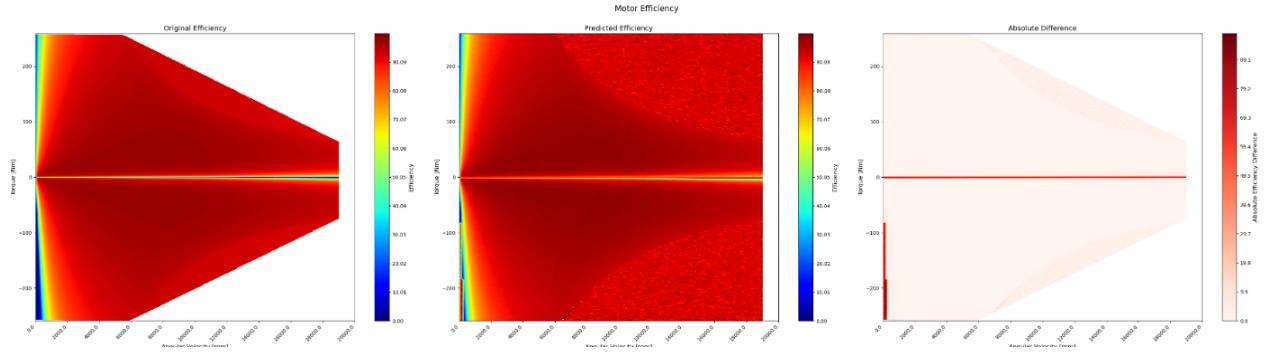


Figure 34: 2nd Baseline Results for 3D KPI(ETA)

We also have the score statistics of how the Baseline model performs over the test dataset.

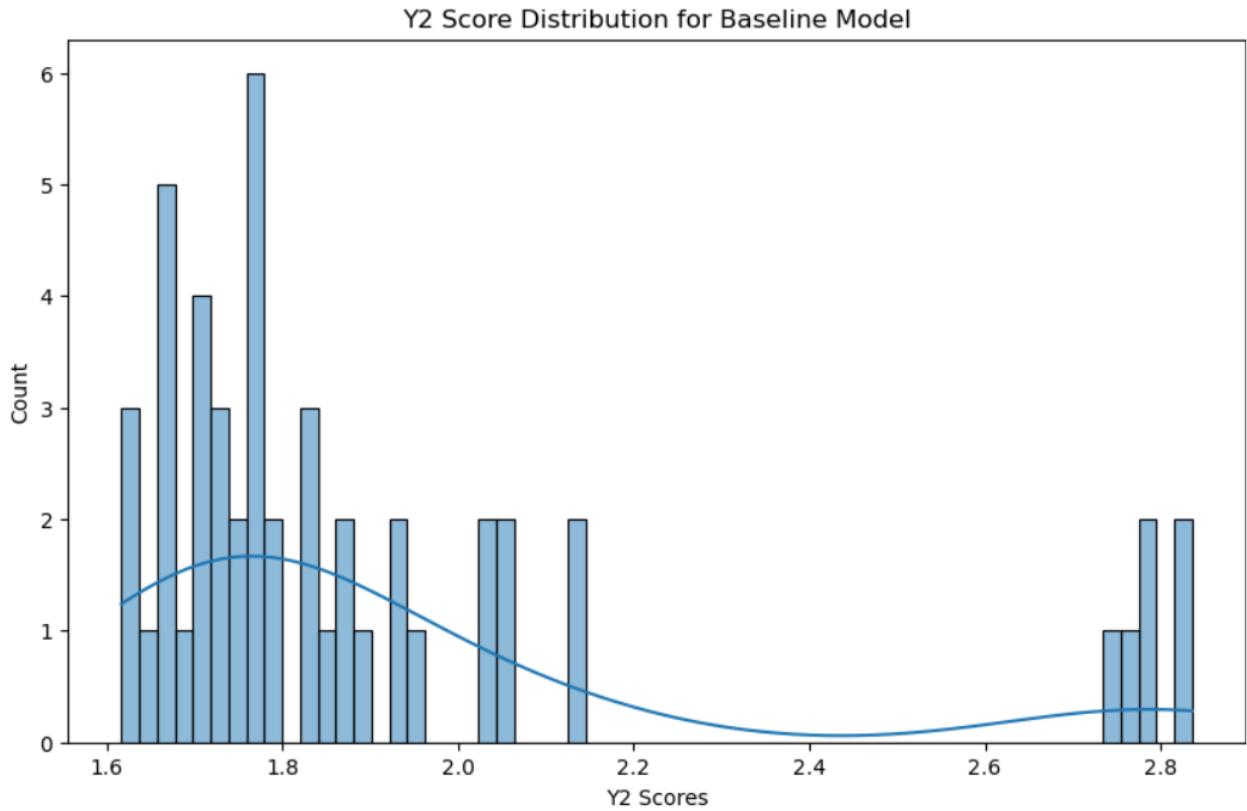


Figure 35: Baseline Score statistics for 3D KPI(ETA)

## Ablation Studies

As part of ablation studies, we have compared our evaluations of both the MLP and the Baseline model on both targets respectively.

<b>Model</b>	<b>Y1 Score</b>	<b>Y2 Score</b>
Baseline	6.170	1.927
MLP	5.1994	0.4962

Table 5: Ablation Studies

We draw the inferences that our model has performed relatively better than the baseline for 2D KPI. The inferences shown are strictly speaking from the Double V Magnet Topology which assumed the bulk of the data we had received.

# Conclusion

## Future Improvements

I would suggest the following improvements to our study : 1. Build a model that uses the prediction from Y1 to predict its corresponding Y2. 2. Although we have designed the model to be topology invariant for 3 topologies, we only had sufficient data from Double V Topology to draw evaluations from. Further evaluations on the other topologies would be beneficial to critically assess our model's performance 3. Additionally, the motivation of building a topology invariant model was the reason we have considered building a heterogeneous graph to model the data. The machinery is elaborated in Section CIE. Such a model could serve as yet another ablation study to our problem.

# List of Figures

1	V1 Magnet (Source: Valeo)	7
2	V2 Magnet (Source: Valeo)	7
3	Nabla Magnet (Source: Valeo)	7
4	EM Design Flowchart	8
5	Complete EM Geometry (Source: Valeo)	10
6	1/8 Motor Crossection	11
7	Standard Deviation of 2D KPI(ETA) Targets	13
8	Standard Deviation of 3D KPI(ETA)	14
9	Standard Deviation of 3D KPI(ETA)	15
10	HetGraph	19
11	Aggregated Training Loss	24
12	Training Loss for Torque Curve	24
13	Training Loss for ETA grid	24
14	Aggregated Training Score	25
15	Training Score for Torque Curve	25
16	Training Score for ETA grid	25
17	Aggregated Validation Loss	25
18	Validation Loss for Torque Curve	25
19	Validation Loss for ETA grid	25
20	Aggregated Validation Score	25
21	Validation Score for Torque Curve	25
22	Validation Score for ETA grid	25
23	MLP Training Results for 2D KPI(Mgrenz)	26
24	Test Dataset RMSE Evaluation for 2D KPI(Mgrenz)	27
25	MLP Score statistics for 2D KPI(Mgrenz)	28
26	1st MLP Training Results for 3D KPI(ETA)	28
27	2nd MLP Training Results for 3D KPI(ETA)	29
28	3rd MLP Training Results for 3D KPI(ETA)	29
29	4th MLP Training Results for 3D KPI(ETA)	30
30	MLP Score statistics for 3D KPI(ETA)	31
31	Baseline Results for 2D KPI(Mgrenz)	32
32	Baseline Score statistics for 2D KPI(Mgrenz)	33
33	1st Baseline Results for 3D KPI(ETA)	33
34	2nd Baseline Results for 3D KPI(ETA)	34
35	Baseline Score statistics for 3D KPI(ETA)	34

# List of Tables

1	Excel File Structure of an EM variant . . . . .	12
2	EM Input Parameters . . . . .	13
3	Hyperparameter Tuning . . . . .	24
4	Scoring Criteria . . . . .	25
5	Ablation Studies . . . . .	35

# Bibliography

# Declaration on oath

I hereby certify that I have written my master thesis independently and have not yet submitted it for examination purposes elsewhere. All sources and aids used are listed, literal and meaningful quotations have been marked as such.

---

Lilly Abraham K64889, 11.12.2024

# Consent to Plagiarism Check

I hereby agree that my submitted work may be sent to PlagScan ([www.plagscan.com](http://www.plagscan.com)) in digital form for the purpose of checking for plagiarism and that it may be temporarily (max. 5 years) stored in the database maintained by PlagScan as well as personal data which are part of this work may be stored there.

Consent is voluntary. Without this consent, the plagiarism check cannot be prevented by removing all personal data and protecting the copyright requirements. Consent to the storage and use of personal data may be revoked at any time by notifying the faculty.

---

Lilly Abraham K64889, 11.12.2024