Concept Proposal by Team Paderborn, Germany

Ravi Pandey
Data Science for Engineering
rpandey@mail.uni-paderborn.de

Oliver Schweins
Machine Learning Applications in Control Engineering
oliverjs@mail.uni-paderborn.de

Nikolas Förster
Power Electronics and Electrical Drives
foerster@lea.uni-paderborn.de

Till Piepenbrock
Power Electronics and Electrical Drives
piepenbrock@lea.uni-paderborn.de

Wilhelm Kirchgässner
Machine Learning Applications in Control Engineering
0000-0001-9490-1843

***Abstract*—In order to comply with the MagNet Challenge 2023 registration criterias, the conceptional approach to the task at hand is described. A mix of black-box, gray-box, and equation-based models is pursued, which shall generate a high power loss estimation accuracy at a very low amount of model parameters.**

***Index Terms*—power magnetics, machine learning, regression, sequence-to-sequence**

I. Introduction

Virtually all electronics depend to some extent on magnetic circuits – be it in consumer hardware, like mobile applications, or industrial power sources. The ever-increasing miniaturization of electronic components and chips on system boards converts thermal and electromagnetic aspects to emerging challenges. A high precision in manufacturing and layout during the design phase becomes more and more urgent, but the mathematical modeling foundation behind power magnetics lacks rigorousness. This renders overdimensioning and large material margins for meeting the minimal requirements robustly an industry standard. In order to overcome these adverse circumstances, the MagNet challenge calls for innovative methods that estimate the power loss in an array of given scenarios consisting of temperature, frequency, and signal waveforms for the magnetic flux and strength in ferrites of a constant geometrical shape. While relative accuracy at a certain model size is one obvious and ambitious goal, another track is denoted by the novelty and explainability of the modeling method, which is expected to give rise to desirable fundamental insights and first principles. Eventually, the user friendliness of the developed modeling tools in the context of sound software engineering is tracked orthogonally to the other two mentioned objectives. With this concept paper, the Paderborn University team sets out its approach and strategy to this challenge, which is to tackle all competition objectives, see Fig. 1.

II. Concept

All solutions will be evaluated on their performance during two phases: first, performance on the ten given materials

This work is supported by the Federal Ministry of Education and Research.

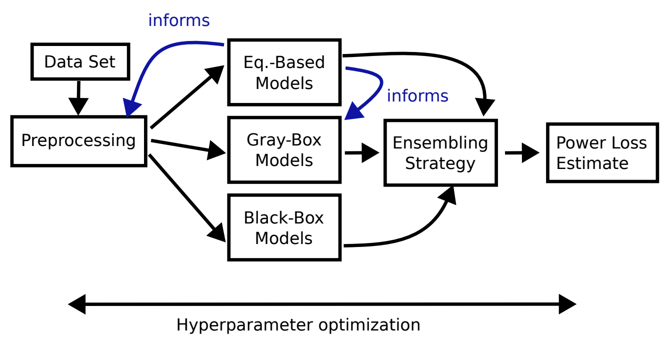

Fig. 1. The trade-off between accuracy with many parameters and high interpretability with few parameters is investigated.

is assessed, and second, the performance on three held-out materials will be taken into account, for which measurement data will be made available only scarcely and at a very late competition stage. These rules are to foster methods that cannot only generalize well to seen materials but also to unseen, new materials. This outline allows for a few different approaches: there can be a (data-driven) model (ensemble) per material, where the learning process on the later three materials has to be as efficient as possible, or there could be just a single ensembling approach, that is able to map materials into an embedded space such that one approach becomes feasible for all materials. The total amount of required model parameters at a high accuracy will most probably decide over which approach is to be followed.

A. Preprocessing

Filters, such as downsampling the signals' 1024 sample points to fewer, as well as the cross-validation (CV) strategy are to be discussed and established upfront. For the latter, a k-fold CV is typical, however, it would certainly require to be stratified such that interpolation during out-of-fold evaluation is encouraged over extrapolation. Moreover, if one model for all materials is seeked for, out-of-fold samples would require to contain observations from other materials, too.

An important aspect is going to be the feature engineering in this competition. Starting with the target, the provided scalar power loss per sample could denote the output of model training (regression), or, more dynamically, the magnetic strength H curve over a time period might be a suitable intermediate estimation target (seq2seq). Ensembling and estimation fusion between the two approaches are obvious. On the model input side, different key characteristics are not provided in the data set upfront. For instance, the signal waveform needs to be classified from the time series, as well as the periodic signal's phase. Furthermore, although a regression model would be capable of processing all 1024 signal samples, it would require plenty of model parameters, whereas several aggregations, e.g., statistical moments, deviations from an ideal shape, etc. might be as reasonable. Eventually, different normalization or transformation schemes could reveal patterns for data-driven models that are easier to discern, and augmentation of the given data through noise application and phase variation might boost generalizability.

B. ML Models

Among data-driven machine learning (ML) models, the following paradigms will be investigated:

- Transformer-based seq2seq models [1],
- Bidirectional GRU/LSTM seq2seq models,
- black-box seq2scalar with neural networks,
- black-box seq2scalar with gradient boosting machines,
- hybrid neural ordinary differential equations seq2seq [2].

The first four methods are black-box, while the latter is gray-box incorporating equations from power magnetics. In contrast to the competition host's tutorial on LSTM models, here a bidirectional approach of LSTM/GRU is followed due to the stationary operation conditions. Moreover, for the regression constellation, gradient boosting machines [3] are a promising paradigm that win many competitions on kaggle. As a method between ML and equation-based paradigms, the hybrid neural ordinary differential equations will be investigated for the sequence2sequence task, where several neural network outputs denote the free parameters of (differential) equations from the power magnetics literature. Such a scheme can account for parameter variability and was demonstrated successfully for thermal modeling in electric machines [4]. Although many methods are followed in an early phase, only the best approaches will be filtered to accommodate the lean model paradigm.

C. Equation-based Models

In addition to the data-driven approaches, equation-based models will also be applied to the given data. Accordingly, some well-known direct loss estimation approaches (Steinmetz, i²GSE, ...) will be fitted and validated with the data. As is usual in practice, these approaches require the preparation of the input signal time series to aggregated quantities for use in the equations. Typical input parameters of these models to predict the power loss density are frequency, signal amplitude and duty cycle, whereas many advanced models require the derivative of the input signal. The derivative can either be calculated approximately (e.g. to a piecewise linear function) or with the full input vector. Different preprocessing methods such as input filters and feature engineering will denote certain hpyerparameters that will be optimized.

Next to the direct loss estimation approaches, also equation-based sequence2sequence models (e.g., Jiles-Atherton, Preisach) will be investigated. These models attempt to estimate the complete BH curve including nonlinearities according to an input waveform. In a second step, the hysteresis losses can be calculated according to the BH curve area. Although, usually not used in practical design approaches, having deterministic sequence2sequence in addition to the established sequence2scalar models may be beneficial in modeling the equivalent data-driven models.

The aim of investigating equation-based models is not to find an independent solution but rather as a base for further (data-driven) investigations. It shall help to understand the effects of different signal properties on the losses. It also aims to provide a benchmark for finding minimum-order material-specific models that achieve the highest modeling accuracy with limited modeling parameters. Furthermore, the limitations of equation-based models in terms of modeling temperature and finding material similarities can be evaluated.

D. Postprocessing

Approximating the scalar power loss by H-curve estimation and area-under-the-polygon-calculation is only up to 5 % accurate. What is more, this deviation is material-specific. Transforming an H-curve estimate material-specific to account for the average deviation from the given power loss is a possible postprocessing scheme to follow. Moreover, fusion of the different model estimates as outlined above will be a task in itself and will probably rely on uncorrelatedness between the models in order to boost overall accuracy.

REFERENCES

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[2] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural Ordinary Differential Equations," in *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018, pp. 6571–6583.

[3] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016.

[4] W. Kirchgässner, O. Wallscheid, and J. Böcker, "Thermal neural networks: Lumped-parameter thermal modeling with state-space machine learning," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105537, 2023.