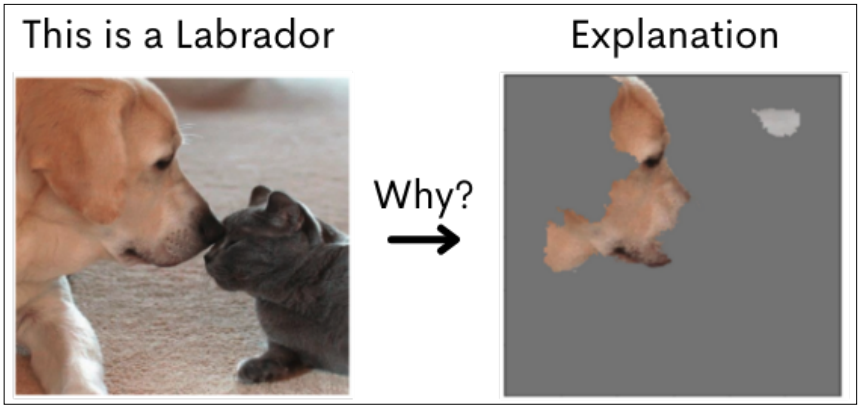
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

DSAR
DATA SCIENCE & ANALYTICS RESEARCH

Master's thesis German or English

Interpretable Deep Learning



About the topic

Nowadays Deep Learning (DL) is a powerful and widely used machine learning (ML) approach with applications in web search, image processing, text analysis, and many more. However, it is often challenging to explain the internal mechanisms of an artificial neural network in a holistic and simple way.

Your task

In our group, we work on novel algorithms to explain decisions of the neural networks. Your main goal will be the investigation of current approaches for deep learning models explanations, including:

- Literature research on new approaches for interpretable deep learning
- Implementation and evaluation of state-of-the-art techniques for the explanation of deep learning models using the python programming language

Your profile

- Backgrounds of computer science, mathematics of other engineering majors
- Excellent knowledge of Python or Java/C++
- Good knowledge of machine learning algorithm is essential.

It is beneficial, if you have visited the "Data Mining & Probabilistic Reasoning" lecture or "Explainable and Fair Machine Learning" seminar by Prof. Kasneci.

An ideal candidate has

- Practical experience working with NumPy, Scikit-Learn, and PyTorch frameworks
- Strong interest in Machine Learning

We offer

- Intensive mentoring
- Research experience for a future career in academics or industry

After the successful completion of the thesis, you will be able to understand the state-of-the-art approaches for explaining the results of artificial neural networks and will know how to implement them using modern machine learning frameworks.

Are you interested?

Please send your CV and current transcript of records to:

Vadim Borisov
Sand 14, C207
vadim.borisov(at)uni-tuebingen.de

Prof. Dr. Gjergji Kasneci
Sand 14, C221
gjergji.kasneci(at)uni-tuebingen.de