3: Analysis of suitable methods for explaining artificial intelligence in the field of autonomous trajectory planning

The progress in the field of autonomous driving is tremendous. New technologies through the development of artificial intelligence are making this progress possible. Today, the first autonomous vehicles are already driving thousands of kilometers on test routes without the need for major intervention of human drivers. Developments in this area are accompanied by increasingly powerful algorithms and methods from the field of machine learning. However, the increasing complexity of the used techniques does not only have advantages, as they are becoming more and more opaque. This intransparency means that certain decisions made by the vehicle are neither recognizable nor understandable to the user, the developer or the legislator. The AI's non-transparent behavior is usually referred to as "black-box" behavior meaning that only the input and output variables are known to the developer. The operation within the black box remains opaque. Explainable AI methods are concerned with resolving precisely this opacity and making complex AI systems understandable and interpretable. However, a problem is that developers and researcher are searching for a quick solution to technical problems, leaving the questions of transparency and accountability on the sidelines. This transparency is necessary for a broad market introduction of autonomous vehicle systems, as it is the basis of trust and legislation implementation.

In the context of this thesis, we will research which methods are suitable and applicable for explaining opaque systems in autonomous driving. The methods and applications will be placed in the overall context of autonomous driving. There are already many methods and algorithms outside of autonomous driving that can be transferred.

The following points are to be worked on:

- Literature research and familiarization with the topic.
- Identification of the requirements for explainability
- Investigation of suitable methods in the area of Explainable AI
- Comparison of methods in the field of Explainable AI
- Classification of the topic in the current state of autonomous driving
- Evaluation and validation of the results
- Comparison to other investigations
- (Optional): Implementation of a small application

You should bring along:

- Creativity
- Independence
- Perseverance
- Interest/knowledge in programming (C++/Python etc.)
- Social competence

The thesis should document the individual work steps in a clear form. The candidate undertakes to complete the Master's thesis independently and to indicate the scientific aids used.

The submitted thesis remains the property of the chair as an examination document and may only be made accessible to third parties with the consent of the chair holder.

Start: immediately   **End: immediately + 6 months**

Chair of Automotive Technology   Supervisor: Rainer Trauth, M.Sc.

Contract: rainer.trauth@tum.de