Title: Machine Learning and Programming Style

Abstract:
Current research here at ULEth is exploring how sociolinguistic differences like gender, region, and experience affect how people use programming languages. Machine learning models are trained with a set of programming samples that are already grouped and labeled. These models can then predict the group the programmer belongs with reasonable accuracy. Alternatively, we can start with a set of samples that we have no groups labels for. In this case the machine learning model will group the samples based on specific features of programming style. We will go over some basic machine learning concepts. Then explore some of the methods being used here at the university, as well as some of the results. Finally, we will discuss some of the future directions of this research and possible applications in the field of software engineering.

The talk should be accessible to anyone. We will define the necessary concepts and examine methods and results more than technical details. Anyone interested in learning how machine learning can be used to solve some problems may gain useful insights.