

## Specialization Module EE

### Analysis and Implementation of Iterative Information Bottleneck and Agglomerative Information Bottleneck

#### Description:

Due to constantly increasing data rates in modern communication systems, new techniques have to be considered to process the data. Since data detection in communication engineering and classification in computer science represent the same problem, techniques from machine learning, pattern recognition, statistical signal processing or artificial intelligence could also be interesting for communication engineering. One step is to cluster the data into a defined number of groups. Since the number of groups is smaller than the cardinality of the original data, clustering can also be seen as compression. If the compression rate falls below the entropy of the data, it is not possible to reconstruct it without loss of information. With the compressed data representation (clusters), the classification / data detection can be done.

One approach for clustering originates from information theory and is called “Information Bottleneck (IB)”. It tries to compress a noisy observation such that it preserves most of a predefined relevant information, i.e. the original source signal. While many of the common clustering techniques use application-specific distortion measures like mean squared error (MSE) or the Euclidean distance, the information bottleneck method is based on the mutual information.

The IB method can be implemented by different algorithms. In this thesis the iterative information bottleneck and agglomerative information bottleneck algorithm have to be implemented. It shall be investigated how the achievable relevant mutual information and rates depend on the number of clusters, how complex both algorithms are as a function of the number of clusters, and how their convergence behavior is.

**Chair:** Prof. Dr.-Ing. Volker Kühn

**Supervisor:** Dipl.-Ing. Clemens-Konrad Müller

**Email:** cm575@uni-rostock.de