Abstract: Third-generation sequencing technologies provide so far unattained read lengths to study genomes and transcriptomes. The direct sensing of unmodified DNA fragments further allows for the detection of base modifications such as 5-methylcytosine (5mC). A characteristic change of the raw nanopore signal within a local sequence context around modified sites is sufficient to discriminate between canonical and modified bases. A widely used base-modification calling algorithm deploys Hidden-Markov Models (HMMs) to test signal segments for the probability to be emitted, given a pore model for native and modified DNA [1]. Artificially induced modifications in GpC contexts have recently been used to combine chromatin accessibility profiling and methylation detection on single molecule level [2].

In this talk I will provide the background for using Hidden-Markov Models as a nanopore raw signal analysis method. Following the crucial steps of signal normalization and alignment, I will introduce a generic model structure based on continuous signal emission probabilities over 6mer contexts. Lastly, I will discuss the relevance of HMMs in the light of modern recurrent neural network approaches with respect to accuracy, training data and computational complexity.

[1]: https://www.nature.com/articles/nmeth.4184
[2]: https://www.nature.com/articles/s41592-020-01000-7

Die Disputation besteht aus dem o. g. Vortrag, danach der Vorstellung der Dissertation einschließlich jeweils anschließenden Aussprachen.

Interessierte werden hiermit herzlich eingeladen

Der Vorsitzende der Promotionskommission
Prof. Dr. K. Reinert