



## Neural Classroom

Recently, Google published WaveNet, a network that is capable of state-of-the-art Text-to-Speech and can also generate speech and music. They achieved generalization across a very diverse corpus by conditioning the network on the identity of a speaker. Furthermore, a network trained on all speakers performed better when generating speech for speaker  $x$  than a network that was only trained on samples from speaker  $x$ . This suggests some kind of transfer learning, which makes sense intuitively. Different people have different ways of speaking and voice characteristics, but there are many similarities as well.

In other work, it was shown that smaller networks could learn to mimic the behavior of much more complex models by training on the activations of the larger network, instead of training on the data directly. The goal of this thesis is to combine these two methods to create a high-capacity teacher network from which smaller student-networks can learn. The motivations for doing this include reduction in storage-space when we do not need the large teacher-network's capabilities, and gaining a better understanding of (transfer) learning. If this sounds interesting to you, do not hesitate to contact us so that we can have a chat.



**Requirements:** Interest in and willingness to study Machine Learning and Deep Learning. There will be weekly meetings to discuss progress and open questions.

**Interested? Please contact us for more details!**

### Contacts

- Gino Brunner: [gino.brunner@tik.ee.ethz.ch](mailto:gino.brunner@tik.ee.ethz.ch), ETZ G63
- Yuyi Wang: [yuyi.wang@tik.ee.ethz.ch](mailto:yuyi.wang@tik.ee.ethz.ch), ETZ 94