Health Information Systems Lab Challenge
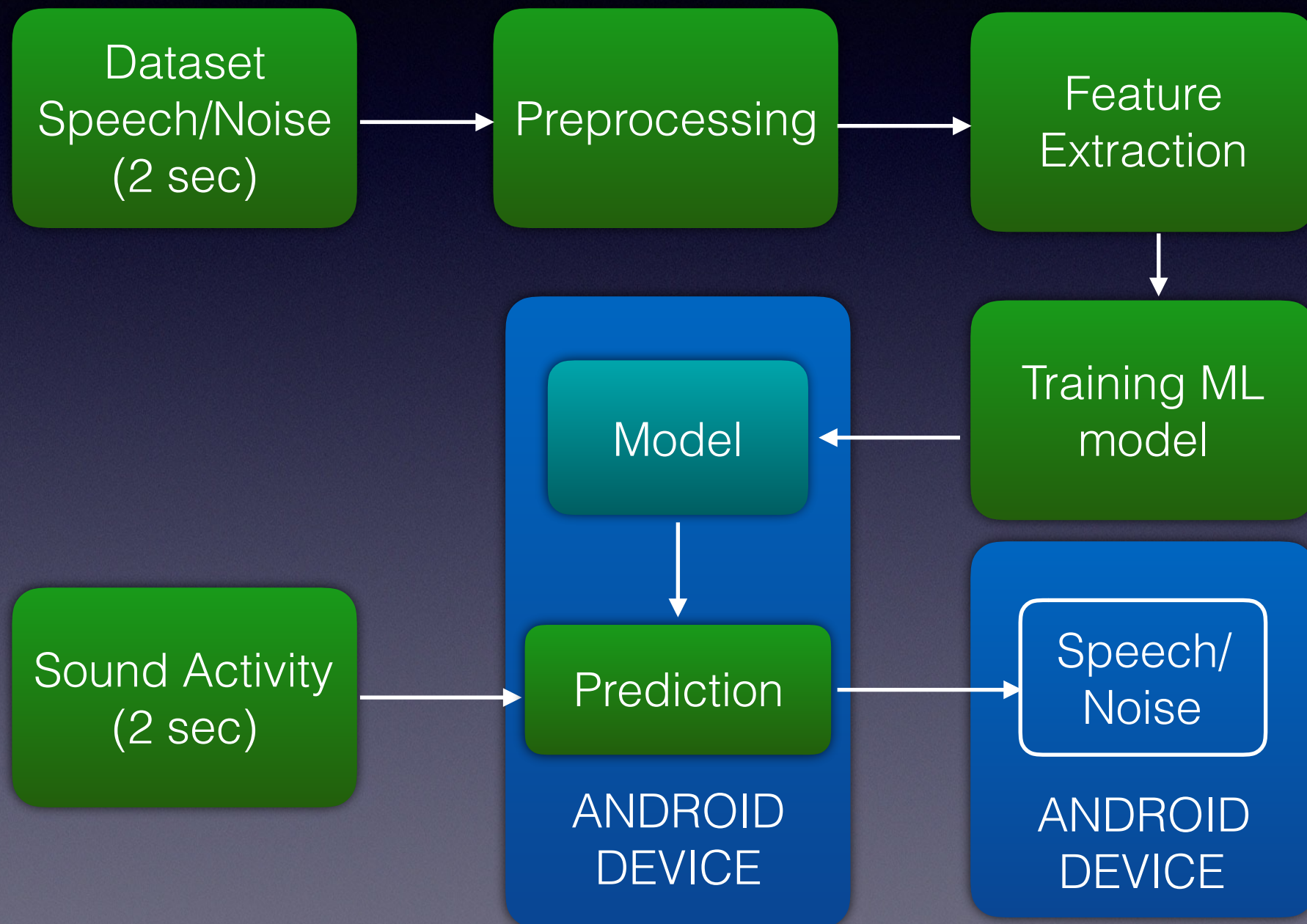
Prabhakaran Santhanam
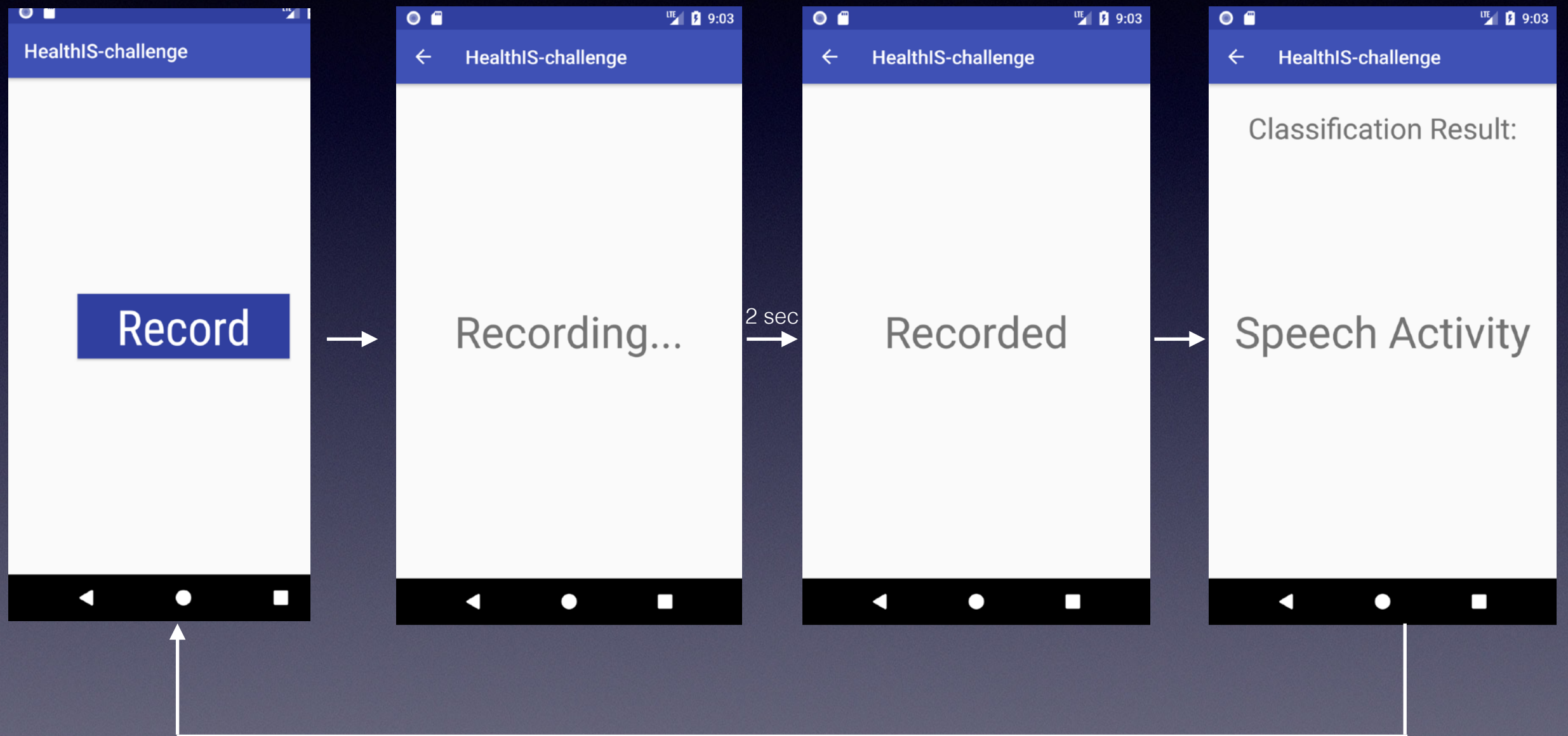MSc Computer Science ETH

Noise/Speech Classification

Dataset

- NOIZEUS: A noisy speech corpus for evaluation of speech enhancement algorithms. <u>Link</u>.

- Each .wav file is for 2 sec. SNR=0 is taken as noise and SNR=15 is taken as speech.

Number of Samples	Total	Speech	Noise
Train	400	200	200
Validation	80	40	40

Model

- Support Vector Machines

- LIBSVM - <u>Link</u>.

- Fine tuning the parameters.

Android UI

Android Implementation

- Permission for writing on external storage.

- Permission to record Audio.

- Load the Model.

- Get the Audio Activity.

- Process the Audio and make prediction.

Future Work

- Dataset: it needs to be large and robust. Then it can be trained with larger models.

- Model: Recurrent Neural Networks, Convolutional Neural Networks for Small-footprint Keyword Spotting and other deep neural networks can be used. Tensor flow can be linked with android for these applications. Link.

- Preprocessing/Feature Extraction: Mel-Frequency Cepstral Coefficients (MFCC), spectrograms can be obtained.

References

- Chang, C.C. and Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), p. 27.

- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H., 2009. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), pp.10-18.

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv: 1603.04467.

- Han, W., Chan, C.F., Choy, C.S. and Pun, K.P., 2006, May. An efficient MFCC extraction method in speech recognition. In Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on (pp. 4-pp). IEEE.

- Vogel, L., 2013. Android development tutorial. Based on Android.