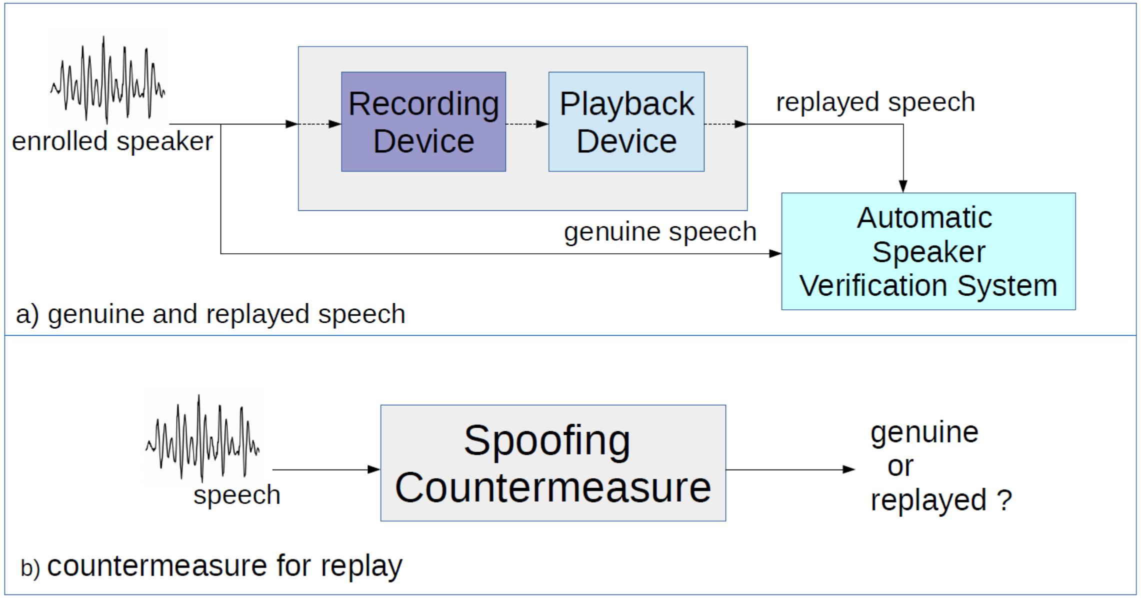
# Analysing the predictions of a CNN-based replay spoofing detection system

Bhusan Chettri

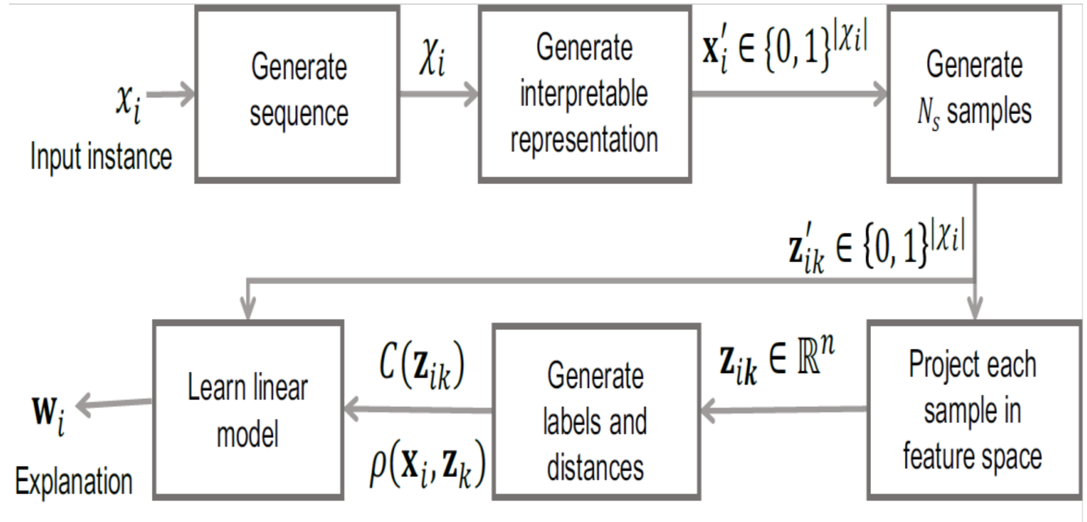
- Understanding the best performing model [2] of the ASVspoof 2017 challenge [1] by generating *temporal* and *spectral* explanations for its predictions using the SLIME [3] algorithm



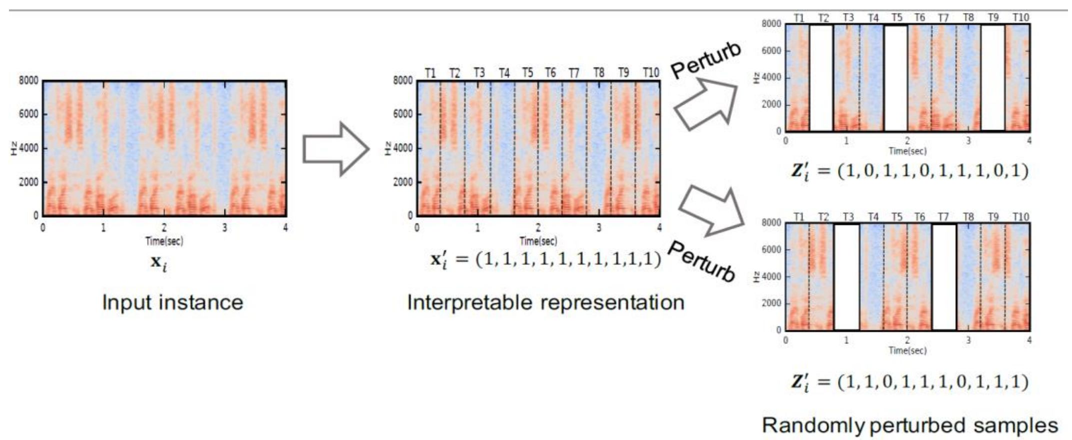

[1] Kinnunen et. al. The ASVspoof 2017 Challenge: Assessing the Limits of Audio Replay Attack Detection in the Wild. In Proc. Interspeech 2017
[2] Lavrentyeva et. al. Audio Replay Attack Detection with Deep Learning Frameworks. In Proc. Interspeech 2017
[3] Mishra et. al. Local Interpretable Model-Agnostic Explanations for Music Content Analysis In ISMIR 2017

# SLIME algorithm

- Top figure: the sequence of steps used by *SLIME* to produce an explanation (in terms of weights) for a given input instance




- Bottom figure: example of segmenting an input instance into 10 uniform temporal components (T1-T10) and generating two samples through random perturbations on these components





[2] Lavrentyeva et. al. Audio Replay Attack Detection with Deep Learning Frameworks. In Proc. Interspeech 2017.
[3] Mishra et. al. Local Interpretable Model-Agnostic Explanations for Music Content Analysis In ISMIR 2017.

# Results and conclusion

- While the model use information across all the frequency bands, more emphasis is given on the first and the last temporal components (T1, T10) for spoofing detection. *We show the significance of our analysis using two interventions*

| | **Dev EER %** | **Eval EER %** |
|---|---|---|
| **I**: Break the system | $7.6 \rightarrow 34.13$ | $10.6 \rightarrow 29.76$ |
| **II**: Protect the system | $7.6 \rightarrow 5.9$ | $10.6 \rightarrow 7.8$ |

- The model gives more importance to the first few milliseconds for class prediction
- Demonstrated the significance of our analysis by preprocessing the test signals that lead to a predictable change in the EER

**Analysing the predictions of a CNN-based replay spoofing detection system**

Bhusan Chettri, Saumitra Mishra, Bob L. Sturm, Emmanouil Benetos

IEEE Spoken Language Technology, Athens, Greece, 18-21 December 2018