Picariello Lectures on Data Science – II Cycle
Università Federico II
February 28 – th, 2022 at 17:00 (CET) on TEAMS

Prof. Vijay K. Gurbani

Project Vāc:
Can a Text-to-Speech Engine Generate Human Sentiments?
Introduces: Prof. Simon Pietro Romano

Sentiment analysis is an important area of natural language processing (NLP) research, and is increasingly performed by machine learning models. Much of the work in this area is concentrated on extracting sentiment from textual data sources. Clearly however, a textual source does not convey the pitch, prosody, or power of the spoken sentiment, making it attractive to extract sentiments from an audio stream. A fundamental prerequisite for sentiment analysis on audio streams is the availability of reliable acoustic representation of sentiment, appropriately labeled. The lack of an existing, large-scale dataset in this form forces researchers to curate audio datasets from a variety of sources, often by manually labeling the audio corpus. However, this approach is inherently subjective. What appears "positive" to one human listener may appear "neutral" to another. Such challenges yield sub-optimal datasets that are often class imbalanced, and the inevitable biases present in the labeling process can permeate these models in problematic ways. To mitigate these disadvantages, we propose the use of a text-to-speech (TTS) engine to generate labeled synthetic voice samples rendered in one of three sentiments: positive, negative, or neutral. The advantage of using a TTS engine is that it can be abstracted as a function that generates an infinite set of labeled samples, on which a sentiment detection model can be trained. We investigate, in particular, the extent to which such training exhibits acceptable accuracy when the induced model is tested on a separate, independent and identically distributed speech source (i.e., the test dataset is not drawn from the same distribution as the training dataset).

Prof. V.K. Gurbani is Associate Professor of Computer Sciences at the Illinois Institute of Technology, co-director of the Real-Time Communication Laboratory and Chief Data Scientist at Vail Systems Inc., where he leads a research group on machine learning, data science and AI activities from idea inception to conceptual proof to prototype and deployment. For a detailed description of his activities look at: http://www.cs.iit.edu/~vgurbani/

The Picariello lectures are jointly organised by the Department of Physics Ettore Pancini & DIETI - Department of Elecrical and Information Technology