POLITECNICO DI TORINO

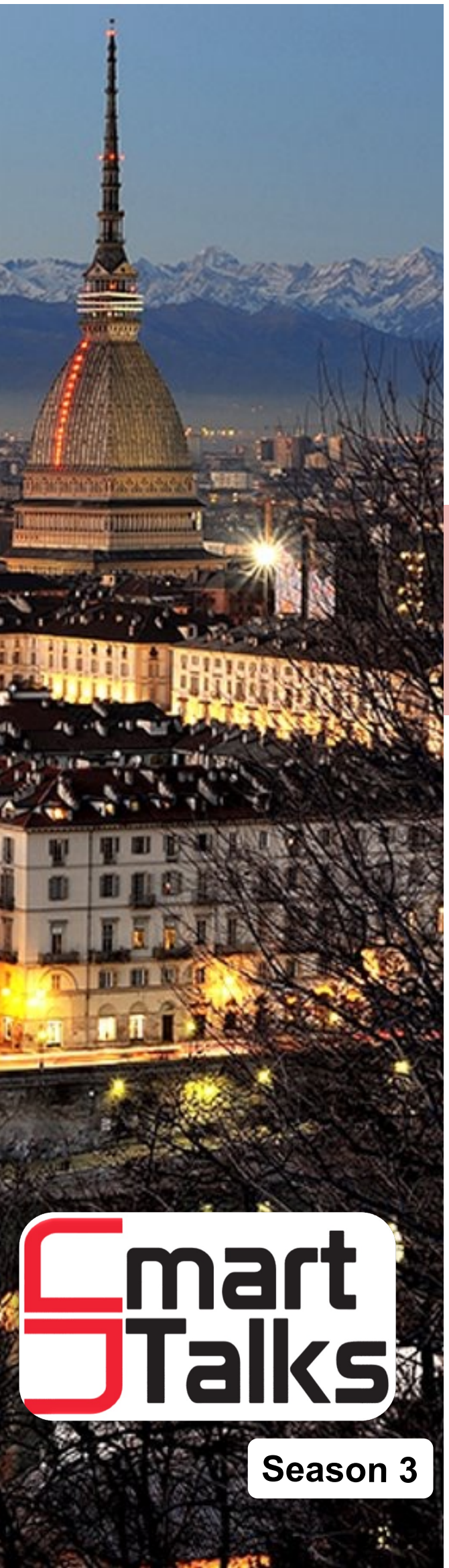
SmartData@PoliTO

SmartData

March 20th 2023, 5:30 PM CEST

SmartTalk: Covivio

https://smartdata.polito.it/category/smarttalks/



Bartolomeo Vacchetti

Politecnico di Torino

Bartolomeo is a PhD student in Creativity-Injection into AI-Powered Multimedia Storyboards at Politecnico di Torino. His research gravitates around the application of machine learning techniques to the image and video field. Specifically, now its main interest lies in discovering and analysing movie editing patterns. He graduated at the Polytechnic of Turin in Movie and Media Engineering. Its main research now gravitates around the discovery and analysis of video editing patterns.

Movie Lens: Discovering and Characterizing Editing Patterns in the Analysis of Short Movie Sequences

ABSTRACT

Video is the most widely used media format. Automating the editing process would impact many areas, from the film industry to social media content. The editing process defines the structure of a video. In this paper, we present a new method to analyse and characterize the structure of 30-second videos. Specifically, we study the video structure in terms of sequences of shots. We investigate what type of relation there is between what is shown in the video and the sequence of shots used to represent it and if it is possible to define editing classes. To this aim, labelled data are needed, but unfortunately they are not available. Hence, it is necessary to develop new data-driven methodologies to address this issue. In this paper we present Movie Lens, a data driven approach to discover and characterize editing patterns in the analysis of short movie sequences. Its approach relies on the exploitation of the Levenshtein distance, the K-Means algorithm, and a Multilayer Perceptron (MLP). Through the Levenshtein distance and the K-Means algorithm we indirectly label 30 seconds long movie shot sequences. Then, we train a Multilayer Perceptron to assess the validity of our approach. Additionally the MLP helps domain experts to assess the semantic concepts encapsulated by the identified clusters. We have taken out data from the Cinescale dataset. We have gathered 23 887 shot sequences from 120 different movies. Each sequence is 30 seconds long. The performance of Movie Lens in terms of accuracy varies (93% - 77%) in relation to the number of classes considered (4-32). We also present a preliminary characterization concerning the identified classes and their relative editing patterns in 16 classes scenario, reaching an overall accuracy of 81%.