SEQUENCE CLASSIFICATION BY RECURRENT VARIANTS OF LEARNING VECTOR QUANTIZATION

J. Ravichandran and T. Villmann
SICIM /Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Learning Vector Quantization (LVQ) methods have been popular choices of classification models ever since its introduction by T. Kohonen in the 90s. Since then, a plethora of improvements have been made over the years to the original formulation of the LVQ algorithm to handle several shortcomings. These days, LVQ is combined with Deep Learning methods to provide powerful yet interpretable machine-learning solutions to some of the most challenging computational problems. However, techniques to model recurrent relationships in the data using prototype methods still remain quite unsophisticated. In particular, we are not aware of any modification of LVQ that allows the input data to have different lengths. This means that they are not particularly suited for applications in domains where the data are sequential, since requiring that all sequences in the dataset be of the same length is severely limiting. Examples of such domains include but are not exclusive to DNA/RNA sequences, time-series data and natural language. To this end, appropriate modifications to the LVQ algorithm are required for sequential data analysis and classification. In this contribution, we propose the use of the Siamese architecture to not only model recurrent relationships within the prototypes but also give the ability to handle prototypes of various dimensions simultaneously.