Learning Vector Quantization with Applications in Neuroimaging and Biomedicine

van Veen, Rick

DOI:
10.33612/diss.211419033

Chapter 7

Summary

The main focus of this thesis is the classification of neurodegenerative diseases using machine learning, particularly learning vector quantization (LVQ).

Neurodegenerative diseases (Chapter 2) are characterized by progressive loss of neurons in a patient's nervous system and may show overlapping clinical symptoms, especially at an early stage. Therefore, [^{18}F] fluorodeoxyglucose (FDG)-positron emission tomography (PET) is used as an ancillary investigation in clinical practice to help neurologists diagnose specific neurodegenerative diseases. FDG-PET measures the uptake of glucose which directly correlates with the activity within each small region of a person's brain. As a diagnostic aid, quantification of activity profiles as measured by FDG-PET would be helpful but is challenging. The use of scaled subprofile model (SSM)/principal component analysis (PCA) has provided promising results when working with a single neurodegenerative disease. However, it has difficulties differentiating between multiple neurodegenerative disorders.

In Chapter 4, we have provided a generalized matrix learning vector quantization (GMLVQ) system that can discriminate between Alzheimer's disease (AD), dementia with Lewy bodies (DLB), healthy control (HC), and Parkinson's disease (PD) patients. Furthermore, we show it is possible to project the high dimensional FDG-PET data onto the three-dimensional decision space of GMLVQ and provide the neurologist with insight into the system's inner workings. Idiopathic rapid eye movement sleep behavior disorder (iRBD) is the earliest definable stage of PD or DLB. We investigate the trajectory towards DLB or PD of iRBD patients by projecting them onto the GMLVQ decision space. This approach, including the suggested improvements in the chapter, might prove to be advantageous in predicting disease progression in early-stage or prodromal patients with neurodegenerative disease.

Obtaining enough FDG-PET data from a single source can be challenging. In Chapter 5, we study machine learning methods to perform an automated classification of HC, AD, and PD subjects' FDG-PET data collected at three neuroimaging centers. Results obtained for the individual centers show that reliable classification is possible. However, we also demonstrate that cross-center performance can decrease problematically due to conceivable center-specific characteristics of the avail-

able FDG-PET data. Additionally, we introduce a technique that uses the linearity of the SSM/PCA method to transform the prototypes and relevance matrix of GMLVQ into the original voxel space of the FDG-PET data. The assessments of three neurologists confirm that prototypes and relevance matrices presented in voxel space are similar to known activity profiles of PD patients.

In Chapter 6, we show that multi-source data leads to systems using sources of variation which are extrinsic to the classes we want to distinguish. Examples of these unwanted sources are different scanners or protocols. Ideally, one would want to get rid of this source of variation by using the same scanners and protocols. However, this is challenging in practice and would not fix currently available data. Therefore, we introduce a new method to deal with the inter-center variation based on GMLVQ. This method is a two-step approach in which a first GMLVQ system is used to find the relevance subspace describing the difference between the centers. In a second step, we use this subspace to correct a GMLVQ system trained on the diagnostic problem, projecting out any center differences. We show that this procedure creates "cleaner" performing models and more expressive relevance profiles on an artificial and multi-center neuroimaging dataset. Finally, the "center" problem is just one example application where a correction is possible. The method applies to other problems as well. For example, in medicine, we often have sex specific variations in the data which overlay the target classification. In a technical context, measurements could come from different platforms or devices.

From a technical perspective, we have presented "sklvq", a Python implementation of generalized learning vector quantization (GLVQ), GMLVQ, and localized generalized matrix learning vector quantization (LGMLVQ). In addition to providing the core functionality for the GLVQ, GMLVQ, and LGMLVQ algorithms, scikit-learning vector quantization (sklvq) is distinctive by emphasizing its modular and customizable design. This design choice (see Chapter 3) results in a feature-rich implementation for users and enables easy extensions of the algorithms for researchers. In the second part of Chapter 3, we discuss a typical use case of GMLVQ in biomedicine. In this study, we have compared the diagnostic performance of GMLVQ based on steroid data obtained using gas chromatography–mass spectrometry (GC-MS) and liquid chromatography–mass spectrometry (LC-MS), intending to classify adrenocortical tumors into their benign and malignant variants. Whether the tumor is malignant or benign plays a critical role in the management of such patients. We found that although fewer steroids can be measured using LC-MS, it has the potential to become the standard versus the currently used GC-MS method, which is much more labor-intensive, time-consuming, and expensive.

7.1 Outlook

This thesis can be extended in several directions. We consider the following topics for future research.

- The data of Chapter 4 can be improved based on the outliers and ambiguous cases pointed out by the GMLVQ models. The included neurodegenerative

diseases contain many sub-variants which have not been included. Three future steps are conceivable. First, GMLVQ could be trained on typical cases only, and sub-variants are later projected in the systems decision space of the system, as is done with the iRBD patients. Second, one could consider the sub-variants as part of the main variant to stretch the decision space over the entire spectrum of a neurodegenerative disorder. Thirdly, if enough data is available, one could consider the sub-variants as separate classes.

- Based on the findings in chapter 5, one could consider looking into different ways of dimensionality reductions that do not use up any of the already "scarcely" available patients. Random projection could be an intriguing candidate.

- The voxel representations (chapter 5) enable the explanation of the decision process of GMLVQ in easier-to-understand formats. One could build a system that generates a text (in any natural language) that explains the difference between the prototypes and a new subject. Furthermore, using the voxel representation of the diagonal of the relevance matrix, such a system could explain why specific differences do or do not matter for the final decision.

- In chapter 6, we concluded that not all "center" variation was removed by the correction procedure. More effort is required to improve the method and compare it to other methods. Specifically for the neuroimaging case, one such method is the idea that healthy controls approximate the center differences, similarly to what a calibration phantom (Hoffman phantom) would do. By centering the data from each location using its healthy controls, the center differences may disappear.

Finally, the results in the fields of neuroimaging and biomedicine are promising. LVQ has shown to be an excellent diagnostic aid both in terms of performance and the ability to reduce the decision space. Furthermore, the methods and results presented in this thesis lay the foundation for many interesting applications and methodological advances.