Autoencoder with Dual Loss Function for follow-up estimation

Marco Dossena[1,*,†], Christopher Irwin[1,†]

[1] *Computer Science Institute (DiSIT), University of Piemonte Orientale, Alessandria, Italy*

Abstract

Autoencoders have emerged as powerful tools for unsupervised representation learning, finding applications in various domains such as computer vision, natural language processing, and anomaly detection. More generally, the latent space reconstruction mechanism can be extended to the reconstruction of any type of data. This extended abstract presents an idea of representation learning applicable to data with an initial instant described by a baseline (input data), and an instant in the future referable to a follow-up (output data). The novel approach consists of combining a construction-focused loss with a classification-driven loss. The proposed hybrid autoencoder architecture aims to simultaneously enhance data reconstruction while learning discriminative features for classification tasks. Initial experimental results demonstrate the efficacy of the proposed hybrid autoencoder on a long-covid dataset.

Keywords

manifold learning, autoencoder, long-COVID syndrome

1. Introduction

Traditional data representation techniques often rely on manual feature engineering, which is labor-intensive, domain-specific, and might miss out on intricate patterns present within the data. Autoencoders [1], a class of neural networks, present an appealing solution to this challenge by enabling the automatic learning of data representations in an unsupervised manner. The core idea of an autoencoder revolves around dimensionality reduction, wherein the network learns a compressed representation of the input data that captures its salient features. This compressed representation, often referred to as ”latent space”, can then be used for various downstream tasks such as classification, reconstruction, and generation. In our research context, the input space corresponds to the patient’s description at the time of hospitalization. In contrast, the output space is expanded to include follow-up data, specifically one year after hospitalization. This augmentation of the output space serves as a valuable means to inform and guide the representation of the latent space within our model. Furthermore, to classify patients who may be suffering from long-covid, we introduce an additional classification head building upon the hidden patient representation.

Doctoral Consortium AIxIA 2023, November 6–7, 2023, Rome, Italy

[*] Corresponding author.

[†] These authors contributed equally.

marco.dossena@uniupo.it (M. Dossena); christopher.irwin@uniupo.it (C. Irwin)

2. Methodology

The autoencoder architecture includes two basic modules: encoder and decoder. In the following, we will present these two components.

Encoder: this module allows the compression of the initial features space into the latent representation. In our setting, the encoder is realized by a fully connected layer that maps the input space into a lower dimensional latent representation. The encoder takes as input a sample $x \in \mathbb{R}^{d_{in}}$ and outputs a representation $h \in \mathbb{R}^{d_{hid}}$, where $d_{in} \gg d_{hid}$. During the encoding phase, we also introduce non-linearity by applying a non-linear activation function immediately after the fully connected layer. Finally, to reduce overfitting and improve the model generalization we also adopt a dropout layer [2].

Decoder: module that reconstructs the output data starting from the latent representation. In most settings, this module is designed to complement the encoder, aiming to reconstruct the input sample x from the latent representation h. However, in our experiments, we chose to expand the output dimension d_{out} to include the features of the patients at follow-up time $d_{out} = d_{in} + d_{fu}$. In this configuration, the encoder needs to map the input features to a latent representation that not only compresses the data but also retains sufficient information for accurate reconstruction of the follow-up details during the decoding phase. As a result, the model can potentially learn patterns that span across the hospitalization and follow-up time.

Lastly, to address the task of classifying patients with or without long-covid syndrome, we incorporate a classification head into the model.

2.1. Dual loss function

During the learning process, the model uses a loss function comprised of two separate components. The first part $\mathscr{L}_r$ is responsible for the reconstruction part of the learning task. In particular, we resorted to the Mean Squared Error (MSE) loss that calculates the distance between the reconstructed data and the training samples. The second component denoted by $\mathscr{L}_c$ is used for the classification of long-covid syndrome, using the Binary Cross Entropy loss (BCE). We chose to mix these two losses by introducing two coefficients, denoted as α and β, enabling us to seamlessly balance between a reconstruction and classification regimen. The complete loss function is described as follows:

$$\mathscr{L} = \alpha\mathscr{L}_c + \beta\mathscr{L}_r$$

3. Experiments

During the experiments, we applied the model to a real-world dataset. In the subsequent sections, we will start by giving an overview of the long-covid scenario, presenting dataset statistics, and finally, we will discuss the model configuration and performances.

3.1. Long-COVID scenario

Following the characterization in [3], long-COVID-19 syndrome consists of signs and symptoms (sequelae) consistent with COVID-19 that are present beyond 12 weeks of the onset of acute COVID-19 infection, and not ascribable to alternative causes (i.e., other diseases). Consider the syndrome to be defined as the persistence of at least one of such symptoms, where instances are the patients' data collected at hospitalization, and the labels are the long-COVID symptoms persisting at follow-up.

Concerning patient characterization, *baseline* data indicate 38 features describing the demographic and medical history of the patient, while *hospitalization* data (14 features) refer to the patient's symptoms at hospitalization (acute COVID-19 onset). Baseline data are not directly related to COVID-19 infection but are important factors to take into account in order to make an accurate diagnosis or prediction. Features in the baseline data can be grouped in terms of demographic characteristics (sex, age, smoking attitude, ...) and of prior comorbidities (obesity, chronic liver disease, hypertension, anxiety and depression, ...).

Hospitalization data include the patient's symptoms at COVID-19 onset (fever, cough, dyspnea, arthralgia, ...), drugs administered (hydroxychloroquine, monoclonal antibodies, glucocorticoids, antivirals, ...), and hospitalization information (duration, oxygen administration, ICU intubation, ..). Baseline and hospitalization jointly form the **input space**.

The *follow-up* data (27 features) contains among others the same symptoms present in the hospitalization but at a different instant (one year in the future). The follow-up set combined with the input space forms the **output space**.

The original dataset consisted of 324 entries, representing a very limited data scenario for a deep learning architecture. To augment our dataset with additional samples, we employed the Synthetic Minority Over-sampling Technique (SMOTE) [4], resulting in an expanded dataset containing over 400 samples.

3.2. Results

We conducted a series of tests on the aforementioned dataset to assess whether the latent space constructed by our model could establish a correlation between hospitalization and follow-up while simultaneously maintaining discriminative capabilities for identifying cases of long-covid. Detailed hyperparameters employed during the model training are provided in Table 1.

To address the limitations posed by the dataset size, we adopted a shallow dimension for the latent space and relatively high dropout probability. This decision is justified by the dataset's size, as maintaining a compact representation enhances the model's generalization capacity when there are limited examples available.

The average accuracy achieved in our experiments is 71% ± 0.9. We applied PCA to the latent space embeddings varying the α parameter of the loss, which governs the classification contribution to the loss. As illustrated in Figure 1, it is clear that as the value of alpha increases, the explained variance of the embeddings also rises. This suggests that the clusters become progressively more linearly separable. The result is especially promising considering the scarcity of data and the complexity of the multi-class classification problem.

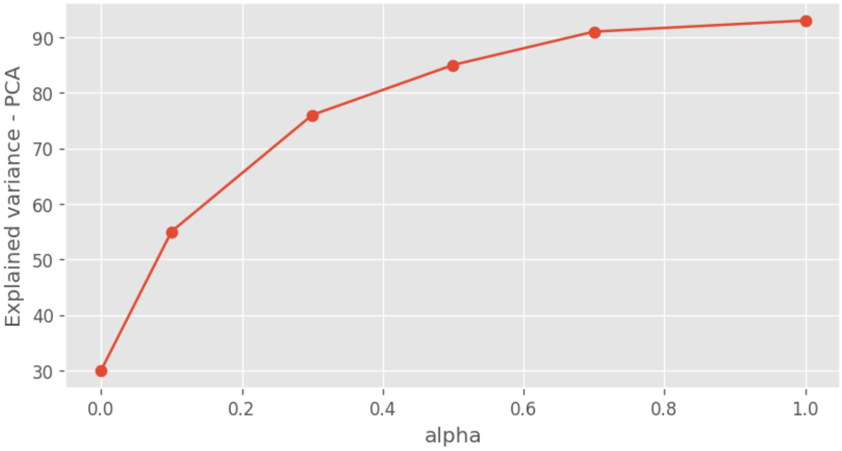

Figure 1: Explained variance of the PCA of the latent space at different values of α parameter of the loss.

3.2.1. Latent space representation.

The dual loss framework enables the creation of a latent space that takes into account the presence of at least one of the symptoms as a discriminator. This capability is made possible by the end-to-end architecture of our model. When visualizing the latent space in two dimensions using t-distributed stochastic neighbor embedding (t-SNE) [5], we observe the emergence of two distinct clusters that represent the distribution of samples within the latent space. Figure 2 illustrates these clusters within the latent space, with sample points color-coded to indicate the presence or absence of symptoms.

Parameter	Configuration
Hidden dimension	128
Latent size	16
α	1
β	0.7
Reconstruction Loss	MSE
Learning rate	0.001
Dropout	0.5

Table 1
The hyperparameters of the model used during the experiments.

4. Conclusion

In conclusion, this paper has introduced a hybrid autoencoder architecture that leverages both construction-focused and classification-driven loss functions to enhance unsupervised

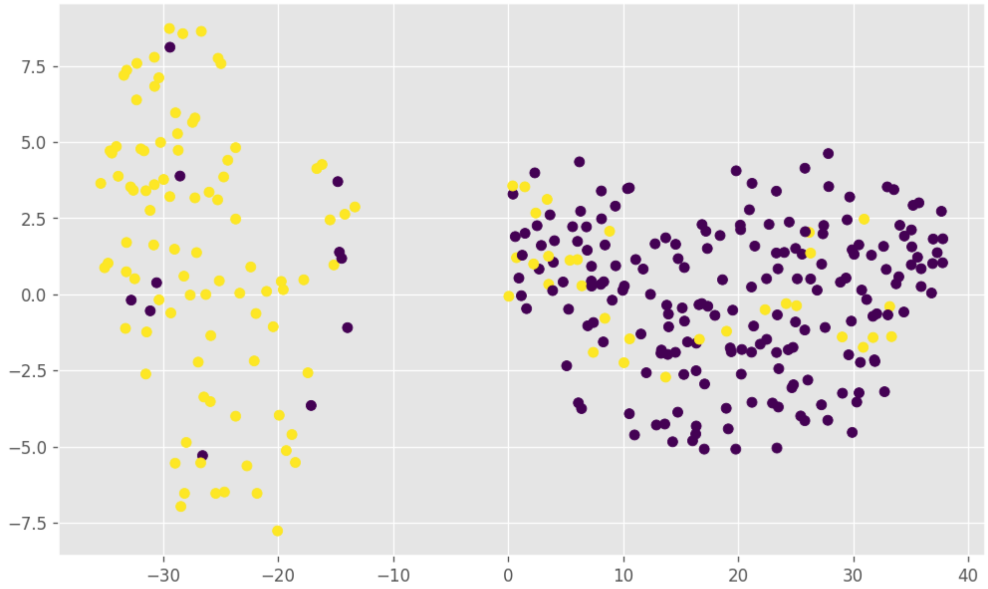

Figure 2: Latent space visualization. Yellow dots represent samples with at least 1 residual symptom at follow-up and vice versa.

representation learning. Autoencoders have already proven their utility in various domains, and this work extends their applicability to data with a temporal aspect. By incorporating both reconstruction and discriminative feature learning objectives, our approach aims to provide a comprehensive solution for a wide range of tasks.
Our initial experiments on a long-covid dataset have yielded promising results, demonstrating the effectiveness of the proposed hybrid autoencoder in capturing meaningful representations from sequential data. These findings pave the way for future research in utilizing autoencoders for time-dependent data and underline the potential impact of this approach in addressing complex real-world problems.

Acknowledgments

M. Dossena and C. Irwin are supported by the National PhD program in Artificial Intelligence for Healthcare and Life Sciences (Campus Bio-medico University of Rome). We want to thank A. Chiocchetti and M. Bellan for having provided us with the long-COVID data and for several fruitful discussions about the case study.
We want to thank our tutors Luigi Portinale, Annalisa chiocchetti, Luca Piovesan and Stafania Montani for their support in our PhD journey.
This work has been supported by the "Piano Riparti Piemonte", Azione n. 173 "INFRA-P. Realizzazione, rafforzamento e ampliamento Infrastrutture di ricerca pubbliche–bando" INFRA-P2-TECHNOMED-HUB n. 378-48 [6].

References

[1] D. E. Rumelhart, J. L. McClelland, Learning Internal Representations by Error Propagation, 1987, pp. 318–362.

[2] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, Journal of Machine Learning Research 15 (2014) 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[3] A. Nalbandian, et al., Post-acute covid-19 syndrome., Nature Medicine 27 (2021) 601–615.

[4] K. W. Bowyer, N. V. Chawla, L. O. Hall, W. P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, CoRR abs/1106.1813 (2011). URL: http://arxiv.org/abs/1106.1813. arXiv:1106.1813.

[5] L. van der Maaten, G. E. Hinton, Visualizing data using t-sne, Journal of Machine Learning Research 9 (2008) 2579–2605. URL: https://api.semanticscholar.org/CorpusID:5855042.

[6] TECNOMED-HUB webpage, 2023. URL: https://www.tecnomedhub.it.