UDC 004.627

DEVELOPMENT OF A VARIATIONAL AUTOENCODER FOR HANDWRITTEN DIGIT RECOGNITION

Podoliak Bohdan, student, b.podolyak.fit.122.20@knute.edu.ua,
State University of Trade and Economics
Filimonova Tetiana, Cand. Sc. (Phys.-Math.), Assoc. Prof,
t.filimonova@knute.edu.ua, State University of Trade and Economics

With each passing day, the application of neural networks becomes increasingly noticeable in various fields of activity. Scientists and researchers strive to develop new and improve existing neural networks to address issues in forecasting, creative tasks, medicine, and particularly image recognition. An essential tool in achieving this goal is variational autoencoders. A Variational Autoencoder (VAE) is a type of neural network applied for encoding data into a latent space. The latent space is a space with lower dimensionality than the data space [1]. These autoencoders, like standard autoencoders, have an encoder and a decoder, but they differ in the way they encode information. Variational autoencoders compress input data to place it in the latent space, which is smaller and more convenient to work with.

Variational autoencoders find wide application in various fields, but they demonstrate the greatest potential in solving image recognition problems, especially in handwritten digit recognition. Handwritten digit recognition is relevant to humanity as it has many practical applications in everyday life, such as automatic mail sorting, check processing, automatic form filling, recognition of various document and credit card numbers, making it a key factor in societal development. The ability of variational autoencoders to work with different types of data (sequential or non-sequential, continuous or discrete, labeled or unlabeled) gives them certain advantages over other methods of handwritten digit recognition [2]. An example is the capability of variational autoencoders to process non-sequential data, such as handwritten digits written in various sizes, styles, and angles. Variational autoencoders make these data more practical and flexible for further applications.

When creating such an autoencoder, the first step is to import all necessary libraries, load data from the dataset, and prepare it (normalize). The next step involves declaring global variables and defining a loss function, which plays a crucial role in the variational autoencoder's training process. Afterward, the variational autoencoder model is created, including defining input images and forming the encoder (Fig. 1) and decoder (Fig. 2) architectures responsible for compressing images into the latent space and generating new images based on the latent representation, respectively.

```
#Encoder Architecture
input_img = Input((28, 28, 1))
x = Flatten()(input_img)
x = Dense(256, activation='relu')(x)
x = dropout_and_batch(x)
x = Dense(128, activation='relu')(x)
x = dropout_and_batch(x)
z_mean = Dense(hidden_dim)(x)
z_log_var = Dense(hidden_dim)(x)
```

Figure 1 – Encoder Architecture

```
#Decoder Architecture
input_dec = Input(shape=(hidden_dim,))
d = Dense(128, activation='relu')(input_dec)
d = dropout_and_batch(d)
d = Dense(256, activation='relu')(input_dec)
d = dropout_and_batch(d)
d = Dense(28*28, activation='sigmoid')(d)
decoded = Reshape((28, 28, 1))(d)
```

Figure 2 – Decoder Architecture

The final step will be adding validation data and training the model. To verify the results, examples of original and reconstructed images were presented (Fig. 3), a distribution plot of vectors in the hidden layer (Fig. 4), and a loss function graph (Fig. 5).

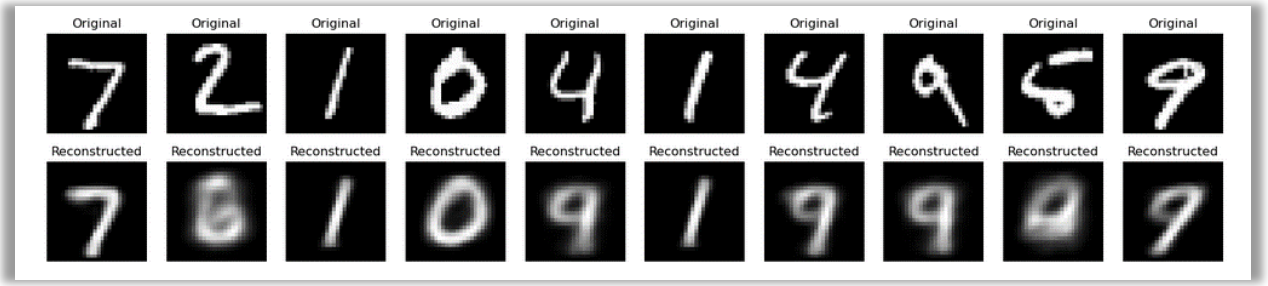

Figure 3 – Autoencoder Operation Visualization

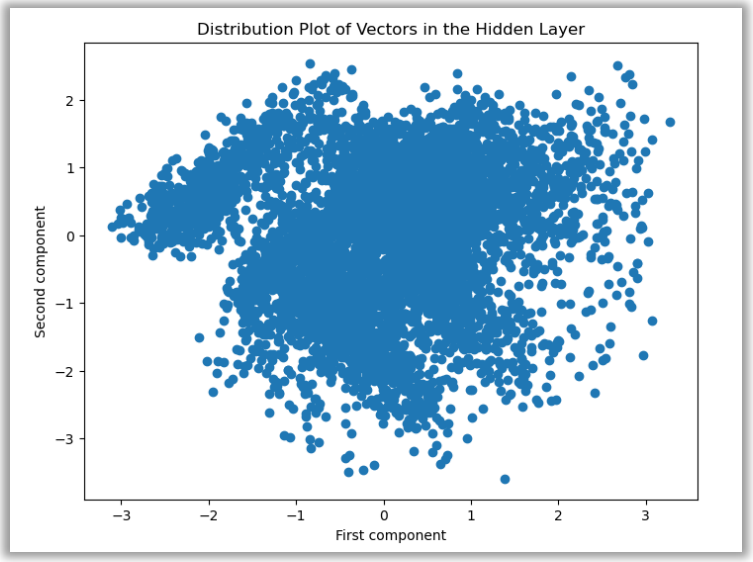

Figure 4 – Distribution Plot of Vectors in the Hidden Layer

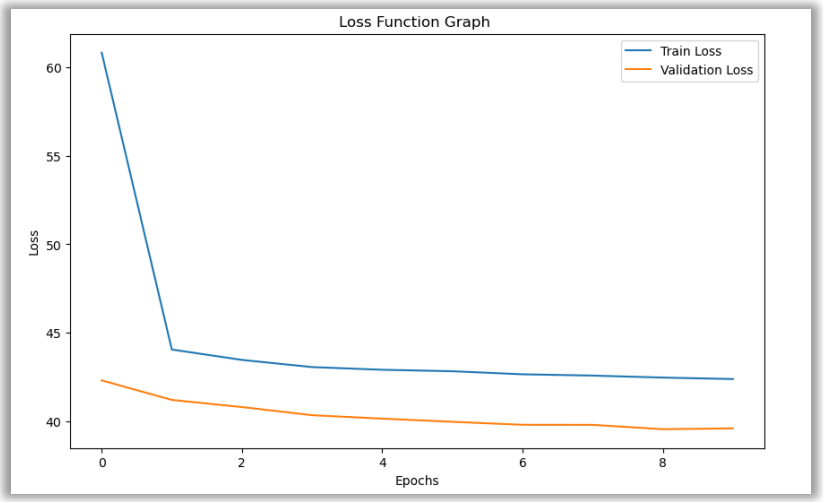

Figure 5 – Loss Function Graph

Conclusion. As a result of the research, the variational autoencoder is quite relevant for handwritten digit recognition and has significant practical potential in other areas. Despite some drawbacks, such as lower speed compared to other methods and the need for a large amount of data for proper functioning, variational autoencoders remain a powerful tool for handwritten digit recognition, with substantial potential for the future.

REFERENCES

1. Variational autoencoder: website. URL: https://en.wikipedia.org/wiki/Variational_autoencoder (Accessed: 14.02.2024)
2. Intuitively Understanding Variational Autoencoders: website. URL: https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf (Accessed: 14.02.2024)