A comparative study for Kuzushiji-MNIST using Self-organizing map and Restricted Boltzmann machine

Rachael Olatunji

Rachael.olatunji@city.ac.uk

Abstract

The objective of this research paper is to provide a comprehensive analysis of two artificial neural networks (ANN) that are trained using unsupervised learning, which in this case would be Self-organizing map (SOM) and Restricted Boltzmann machine (RBM). Both algorithms are employed on a Kuzushiji-MNIST dataset. It was concluded that SOM model showed better results.

1. Introduction

The Kuzushiji-MNIST (KMNIST) dataset comprises of images from the 10 main Japanese hiragana character groups. Hiragana was the first Japanese alphabet [1]. Each letter is unique with the appearance of curved lines. The Hiragana alphabet goes all the way back to the 9th century and is regarded as the basic alphabet used a great deal in literature across centuries [2]. The ability to be able to classify KMNIST would help contribute to the collaborative work in the field of natural language processing (NLP), which could help break down the ambiguity in various languages and be applied in areas such as speech recognition and text analytics [3].

This paper aims to investigate which out of the two models for classifying Kuzushiji-MNIST is more superior based on how accurate the results are. The models that would be intently investigated are Self-organizing map (SOM) and Restricted Boltzmann machine (RBM). For both models, the hyperparameters will be adjusted to see if that influences the accuracy and by how much.

The dataset that would be worked upon for training and testing the model is depicted in brief for section 2. Section 3 evaluates the methodology used by both models during the stage of implementation. Section 4 assesses the outcome from the results whilst providing critical feedback from evaluating the model. The conclusion would be shown in section 5.

1.1 Self-organizing map (SOM)

A self-organizing map (SOM) or self-organizing feature map (SOFM) is an unsupervised machine learning technique that generates a low-dimensional (usually two-dimensional) representation of a higher-dimensional data set while maintaining the data's topological structure [4]. It is a type of artificial neural network (ANN) [5]. A self-organizing map (SOM) is a grid of neurons that conforms to the topology of a dataset, enabling it to explore huge volumes of data and find possible clusters. Through constantly shifting its neurons closer to the data points the SOM will learn the structure of that dataset [6]. SOM is taught using unsupervised learning, which differs from other artificial neural networks in that it does not learn by backpropagation with SGD; instead, it uses competitive learning to alter weights in neurons [7]. Figure [1] below shows the layout of a self-organised map.

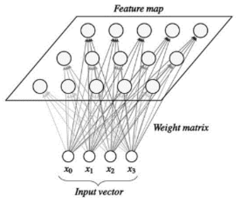


Figure [1]. A figure of the layout of SOM

The main measurements used to judge the quality of SOM are quantization error and topographical error. The average difference between the input samples and the winning neurons is known as quantization error (BMU). It evaluates the accuracy of the original data set, therefore the lower the value, the better [8].

1.2 Restricted Boltzmann machine (RBM)

Restricted Boltzmann machine (RBM) invented by Geoffrey Hinton, is a classification, collaborative filtering, regression, dimensionality reduction, feature learning and topic modelling algorithm [9]. It is a form of artificial neural network (ANN) and is used in probability distribution machine learning. Artificial neural network is a hardware and/or software system that mimics the way neurons in the human brain works [10]. Restricted Boltzmann machines (RBM) are probabilistic nonlinear feature learners that are unsupervised [11]. The RBM architecture is divided into two layers: hidden and visible, without any dependencies between them. An example of this framework is in Figure [2] with the hidden nodes indicated by grey circles and the visible nodes outlined by white circles [12].

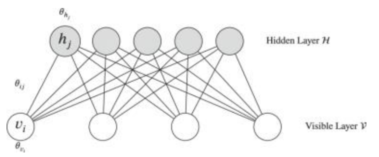


Figure [2]. Layout of RBM

They can automatically find intrinsic patterns in data by reconstructing the input [13]. An RBM converts the inputs into a series of numbers that encodes the inputs in the forward pass. These set of numbers is then translated back into reconstructed inputs in the backward pass [14].

2. Dataset

The Kuzushiji-MNIST (KMNIST) dataset is the more advanced version of the MNIST dataset, where one character is chosen to represent every one of the 10 rows of Hiragana whilst also building up the Kuzushiji-MNIST since MNIST is limited to ten classes. The KMNIST (Kuzushiji-MNIST) dataset contains a training set of 60,000 samples and a testing set of 10,000 examples of handwritten Kuzushiji (cursive Japanese) Hiragana letters. The handwritten characters were resized to fit into grayscale photos at a resolution of 28 x 28 pixels. The KMNIST dataset is recommended for data scientists who want to experiment with machine learning and computer vision on real-world data with little preparation and formatting. The dataset is called by opening Kuzushiji-MNIST as shown below.

```
Kmnist = fetch_openml('Kuzushiji-MNIST', cache=False)
```

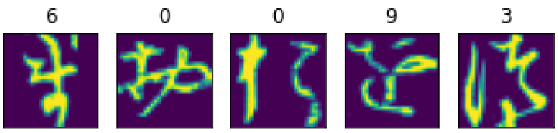


Figure [3]. Samples of a few of the letters in the dataset.

2.1 Initial Data analysis

The objective of such graphic display as part of the initial analysis of the data is to enable the effort to acquire some specific information concealed in the dataset.

Figure [4] shows how each number is classed to represent the letters from the Kuzushiji-MNIST dataset.

```
{0: 'お', 1: 'き', 2: 'す', 3: 'つ', 4: 'な', 5: 'は', 6: 'ま', 7: 'や', 8: 'れ', 9: 'ん'}
```

Figure [4]. What each class represents

Figure [5] shows how much of the count is in each class. A quick observation would point out how they are roughly equal amounts of data in each class, but not exact. This small imbalance would not have a great impact on the training, meaning that the algorithm when trained will not be biased towards a particular class.

	Label	Count
0	0	5261
1	1	5263
2	2	5246
3	3	5232
4	4	5297
5	5	5157
6	6	5280
7	7	5226
8	8	5298
9	9	5240

Figure [5]. Table showing the count for each class

Shown in figure [6] is the table in figure [5] visually represented into a bar graph. As you can see label 5 has a considerably less data than the rest, so the data would be more likely to ignore that class during training.

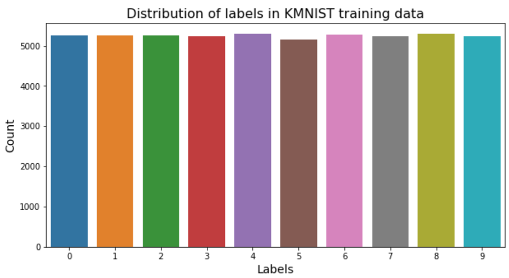


Figure [6]. Bar chart visually representing the count for each class

3. Methods

We describe the structure and hyperparameters used to generate the SOM and RBM models in this part, including how the training, validation, and testing stages were carried out.

3.1 Methodology

The methodology entails reserving 10000 samples from the original dataset as test data for comparing the best SOM and RBM models. The remaining data (60000) is used for training the model selection phase, as well as for training in the algorithm comparison process.

Through competing for representation, every data point inside the data collection knows itself. Initializing the weight vectors is the first stage in the SOM mapping process. Next, at random, a sample vector is chosen, and the map of weight vectors is explored for the weight that best describes that sample. Each weight vector has weights in its immediate vicinity. The chosen weight is rewarded by the ability to become increasingly like the randomly picked sample vector. The neighbours of that weight are rewarded, as they can become more like the selected sample vector. This enables the map to expand and take on new forms. In 2D feature space, they usually form square/rectangular/hexagonal/L shapes. The weights of each node are set. From the training data set, a vector is picked at random. Each node's weights are compared to the input vector to determine which is the most similar. The Best Matching Unit is the name given to the winning node (BMU). The BMU's neighbourhood is then computed. Over time, the number of neighbours reduces the winning weight is acknowledged by resembling the sample vector more. The neighbours start to resemble the sample vector as well. The nearest a node is to the BMU, the more its weights change, and the farther a neighbour is from the BMU, the less it learns.

This procedure would be repeated N times [15].

By rebuilding the input, RBMs discover patterns and extract essential features in the data. As a result, the learning process is divided into numerous forwards and backward passes in which the RBM attempts to recreate the input data. The neural net's weights are changed so that the RBM can detect associations between input features and afterwards identify which ones are significant. Following training, the net may rebuild the input using what it has learned [16].

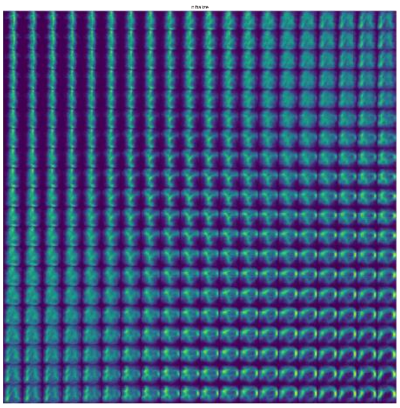
3.1 Hypothesis statement

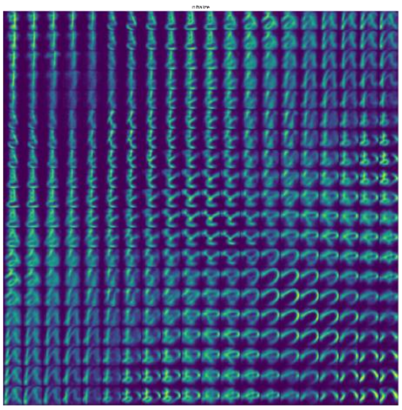
Through conducting literature research from papers based on image classifications using the models SOM and RBM, it is hypothesised that SOM would show better results in comparison to RBM. SOM would also produce quicker results with less need of tuning so easier to run.

3.2 Evaluation methodology

The primary benefit in employing SOM is that the data is simple to read and comprehend. Because of this, the reduced dimensionality and grid clustering, makes it easy to spot patterns in the data. The disadvantage of applying SOM model would be It doesn't establish a generative model for the data, which means it doesn't comprehend how data is generated. It is not as kind with categorical data, and it is considerably worse with mixed categories data. It takes a long time to prepare a model, and it's difficult to train against data that changes slowly. The perks of using RBM is that due to the limitations in terms of connections between nodes, it is faster than a typical Boltzmann Machine. A drawback to RBM would be the weight adjustment in RBM is not as precise as it is in the backpropagation method.

4 Results, Findings and Evaluation





```
It took 111.01423501968384 seconds
```

The initialization of the training data took over 111 seconds.

When the average distance is large, the neighbouring weights are dissimilar, and the weight's location is given a light colour. A darker hue is assigned if the average distance is low. The generated maps reveal that ten zones have higher concentrations of distinct species clusters. The first graph solely shows where species density is higher (darker regions) or lower (lighter regions) (lighter regions). The second illustration shows how they are grouped together.

The RBM did not train as well as I would have hoped. It was very difficult to identify the letters in the graph and how well they were grouped together. It would have been better if the hyperparameter was tuned to get the optimal results with the highest accuracy.

5 Conclusion

The goal was to compare two image classification models using an MNIST dataset. Training and classification are the two primary stages of this type of neural model. In future study, parameters like epochs will be defined in the model in order to collect results like the confusion matrix. Furthermore, more training will be done to optimise the output and acquire value for the accuracy.

6 Reference

[1]. KCP International. 2022. Hiragana: The First Japanese Alphabet | KCP International. [online] Available at: <https://www.kcpinternational.com/2011/06/hiragana-the-first-japanese-alphabet/#:~:text=Hiragana%20are%20usually%20arranged%20in,for%20Si%20doesn't%20exist

[2]. Flexi Classes Japanese. 2022. Japanese Alphabet {Hiragana, Katakana & Kanji} // Why Are There Three?. [online] Available at: <https://flexiclasses.com/japanese/alphabet/#hiragana

[3]. Sas.com. 2022. What is Natural Language Processing (NLP)?. [online] Available at: <https://www.sas.com/en_gb/insights/analytics/what-is-natural-language-processing-nlp.html#:~:text=NLP%20is%20important%20because%20it,speech%20recognition%20or%20text%20analytics.

[4]. En.wikipedia.org. 2022. Self-organizing map - Wikipedia. [online] Available at: <https://en.wikipedia.org/wiki/Self-organizing_map#:~:text=A%20self%2Dorganizing%20map%20(SOM,topological%20structure%20of%20the%20data.> [Accessed 15 May 2022].

[5]. Medium. 2022. Self Organizing Maps. [online] Available at: <https://medium.com/@abhinavr8/self-organizing-maps-ff5853a118d4#:~:text=A%20self%2Dorganizing%20map%20(SOM,method%20to%20do%20dimensionality%20reduction

[6]. Algobeans. 2022. Self-Organizing Maps Tutorial. [online] Available at: <https://algobeans.com/2017/11/02/self-organizing-map/>

[7]. Medium. 2022. Self Organizing Maps. [online] Available at: <https://towardsdatascience.com/self-organizing-maps-1b7d2a84e065

[8]. https://www.intechopen.com/chapters/69305

[9]. 2022. [online] Available at: <https://wiki.pathmind.com/restricted-boltzmann-machine

[10]. WhatIs.com. 2022. What is restricted Boltzmann machine (RBM)? - Definition from WhatIs.com. [online] Available at: <https://www.techtarget.com/whatis/definition/restricted-Boltzmann-machine-RBM

[11]. scikit-learn. 2022. 2.9. Neural network models (unsupervised). [online] Available at: <http://scikit-learn.org/stable/modules/neural_networks_unsupervised.html

[12]. 2022. [online] Available at: <https://www.researchgate.net/figure/An-example-restricted-Boltzmann-machine-RBM-consisting-of-two-layers-one-hidden_fig2_328894173

[13]. Latentview.com. 2022. [online] Available at: <https://www.latentview.com/blog/restricted-boltzmann-machine-and-its-application/#:~:text=Restricted%20Boltzmann%20Machines%20are%20stochastic,two%20layers%20visible%20and%20hidden.

[14]. Medium. 2022. An Intuitive Introduction Of Restricted Boltzmann Machine (RBM). [online] Available at: <https://medium.datadriveninvestor.com/an-intuitive-introduction-of-restricted-boltzmann-machine-rbm-14f4382a0dbb

[15]. Medium. 2022. Self Organizing Maps. [online] Available at: <https://medium.com/@abhinavr8/self-organizing-maps-ff5853a118d4

[16]. Medium. 2022. Reconstruction Of Images Using RBM. [online] Available at: <https://medium.datadriveninvestor.com/dimensionality-reduction-and-feature-extraction-with-rbm-f499965979e9

Figure [1]. Anon, Google image result for https://miro.medium.com/max/606/1*1gooozouknxkzuae9mpzwg.png. Available at: https://www.google.com/imgres?imgurl=https%3A%2F%2Fmiro.medium.com%2Fmax%2F606%2F1%2A1gooOZOUKnXKzuae9mPZWg.png&imgrefurl=https%3A%2F%2Fmedium.com%2Fmachine-learning-researcher%2Fself-organizing-map-som-c296561e2117&tbnid=uXcNKUw6JmjfDM&vet=12ahUKEwja-5-iyt33AhXxVvEDHQVzBQkQMygAegQIARB6..i&docid=HGtUoGoRxqXlfM&w=606&h=451&q=how+does+som+ml+work&hl=en&ved=2ahUKEwja-5-iyt33AhXxVvEDHQVzBQkQMygAegQIARB6

Figure [2]. Anon, Summary arxiv:1612.01158v3 [stat.ML] 25 sep 2018 - researchgate.net. Available at: https://www.researchgate.net/profile/Andee-Kaplan/publication/328894173_Properties_and_Bayesian_fitting_of_restricted_Boltzmann_machines/links/5d82458b299bf1996f7597f6/Properties-and-Bayesian-fitting-of-restricted-Boltzmann-machines.pdf

Glossary

SOM - Collection of neurons arranged in a 1 or 2-dimensional array, with weight vectors corresponding to points in an n-dimensional feature space for all neurons

RBM - A form of artificial neural network (ANN) used to learn probability distributions by machine.

Bernoulli - The discrete probability distribution of a random variable that has a chance of taking the value 1 and a chance of taking the value 0.

Artificial neural network (ANN) – A neural network (artificial neuron network) is a computational model that simulates the way nerve cells in the brain work.

Implementation details

```
from sklearn.datasets import fetch_openml
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
from torchvision import transforms
```

Call KMNIST to open the dataset.