UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Adilet Dossymbekov

Image similarity usage in order to find similar items in E-commerce dataset

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor(s):
Engineer in Computer Vision, PhD, Egils Avots
Science and Technology, BSc Aleksandr Makarov

Tartu 2024

Image similarity usage in order to find similar items in the E-commerce dataset.

Abstract:

This thesis looks into the e-commerce situation among small businesses and examines the possibility of the application of neural networks to business-to-consumer interactions. The goal of the study is to test the ability of NN to work with limited data and hardware.

Pre-trained ResNet-50 model was used as a base, and fine-tuned with the 4 class dataset, consisting of 2906 images of clothes categorised into four classes. The experiments were conducted on a budget-friendly hardware setup. Both the dataset and hardware reflect the limited resources of the small and micro businesses

The final model was evaluated, reaching an accuracy of 70-80%, These results suggest the possibility of NN usage in limited data and computational resource conditions and provide the base for further developments in implementing NN among small businesses.

Keywords:

E-commerce, image classification, neural network

CERCS: P160,P176

Pildi sarnasuse kasutamine sarnaste esemete leidmiseks e-kaubanduse andmestikust.

Lühikokkuvõte:

See bakalaureusetöö uurib väikeste ettevõtete e-kaubanduse olukorda ja hindab võimalust rakendada närvivõrke ärilt tarbijale suhtluses. Uuringu eesmärk on testida närvivõrkude (NN) võimekust töötada piiratud andmete ja riistvaraga. Uuringu aluseks kasutati eelnevalt treenitud ResNet-50 mudelit, mida täiendati neljaklassilise andmestikuga, mis koosnes 2906 rõivapildist, mis olid jaotatud nelja klassi. Katsed viidi läbi soodsa hinnaga riistvaraseadistusega, mis peegeldab väikeste ja mikroettevõtete piiratud ressursse.

Lõplikku mudelit hinnati ja saavutati täpsus 70-80%. Need tulemused viitavad võimalusele kasutada närvivõrke piiratud andmete ja arvutusressursside tingimustes ning pakuvad alust edasisteks arendusteks närvivõrkude rakendamisel väikeste ettevõtete seas.

Märksõnad:

E-kaubandus, pildiklassifikatsioon, närvivõrk

CERCS: P160, P176

(Tranlslated by OpenAI (2024) Chat GPT)

TERMS, ABBREVIATIONS AND NOTATIONS

E-commerce is a commercial activity conducted via electronic media, esp. on the Internet. Although, the source refers to the sector of the economy engaged in such activity. The word itself can be used as an adjective in combination with business terminology such as market, activity, transactions, etc.

B2B - Business-to-Business

B2C - Business-to-Consumer

PwC - PricewaterhouseCoopers International Limited is a multinational professional services brand of firms, operating as partnerships under the PwC brand. It is the second-largest professional services network in the world and is considered one of the Big Four accounting firms, along with Deloitte, EY, and KPMG.

NN - Neural Network

CNN - Convolutional Neural Network

RGB - Red, Green, Blue

ReLU - Rectified Linear Unit

ResNet - Residual Network, a type of architecture of NN introduced in 2015

INTRODUCTION

Rapidly digitalising, and globalising world induced the development of e-commerce activity all around the globe. Further increases caused by the COVID-19 pandemic led to the rise in the e-commerce retailing market, and widening its abilities. Small businesses arise and adapt to developing markets [1], [2].

At the same time, Neural Networks are continuing to develop rapidly, and large businesses such as Amazon are already implementing them [3]. These technologies offer the potential to enhance business-to-consumer interactions by providing advanced image-based search and recommendation systems.

This thesis explores the feasibility of implementing neural networks in small business e-commerce retailing sector settings to improve their product search capabilities. The focus is on evaluating the ability of CNNs to handle limited data and computational resources typical of small and micro-businesses.

This study aims to demonstrate the possibility of using CNNs to enhance e-commerce retailing capabilities for small businesses, providing a foundation for further research and development in this area. By leveraging the power of neural networks, small businesses can potentially improve their competitive edge in the digital marketplace

1 LITERATURE REVIEW

1.1 E-commerce

1.1.1 Early history of E-commerce

The field of E-commerce started its development in the 1970s and the first E-commerce activity was mainly bonded with Electronic Data Interchange technology, which allowed the exchange of data between businesses via phone cables. E-commerce is divided into two types of activities, based on its participants: business-to-business (B2B) and business-to-consumer (B2C). Despite being more recognisable in public B2C types of e-commerce due to its participants (Alibaba, Amazon, etc.), B2B is prevalent in comparison to B2C commerce in terms of revenue.

Further development of e-commerce is correlated with the evolution of IT (Informational Technology). With the Launch of the World Wide Web in 1990, the expansion of the potential client base happened online, so B2C commerce activity started its growth. Usually, those were simple interactions between traders and consumers [4].

For instance, the first online secure transaction was achieved and reported in 1994. Showing that the Internet has become commercially active [4]. A year later, a new significant step in the history of e-commerce was made with the launch of Amazon.com and eBay. The format and content of both sites quickly developed into a fully functional online market, and their success has influenced an increase in e-commerce.

At the same time, the so-called ‘dot-com boom’ or basically, the economic bubble of e-commerce and other related companies started to grow. Despite its collapse, which ended by October 2002, large e-commerce companies survived (Amazon, Cisco etc.). Since then B2C e-commerce activity such as retail e-commerce sales, kept growing yearly [5].

1.1.2 E-commerce nowadays, global pandemic influence

Despite the quantitative growth of e-commerce. Its shape also metamorphosed with the development and change of the Internet environment. Part of the B2C e-commerce activity transferred to mobile devices since consumers started using them for online shopping. Another instance is social media such as Facebook or Instagram, which are new channels for businesses to interact with customers [5].

Figure 1. Forbes Advisor statistics about e-commerce [1].

The COVID pandemic affected all aspects of life, and e-commerce met it with a rapid increase in its market. Since the start of the pandemic physical businesses were forced to close and customers sought the products online. So the composition of e-commerce products drastically changed. Although changes were positive the amount of revenue was expected to meet the plank of 6.5 trillion dollars by 2023 [6], which also cross-confirmed with statistics of Forbes Advisor (Figure 1).

Even before the pandemic, experts predicted the e-commerce retailing sector would outnumber the physical sector in revenue in 2024 only in the USA [7].

The same trends are observable in Kazakhstan, PwC analysis of the first half of 2023 (October 2023). According to their studies, which they believe represent 85 per cent of all retail e-commerce markets in Kazakhstan, sales volumes in the first half of 2023 increased by 79 per cent, compared to the same period of the previous year (Figure 2).

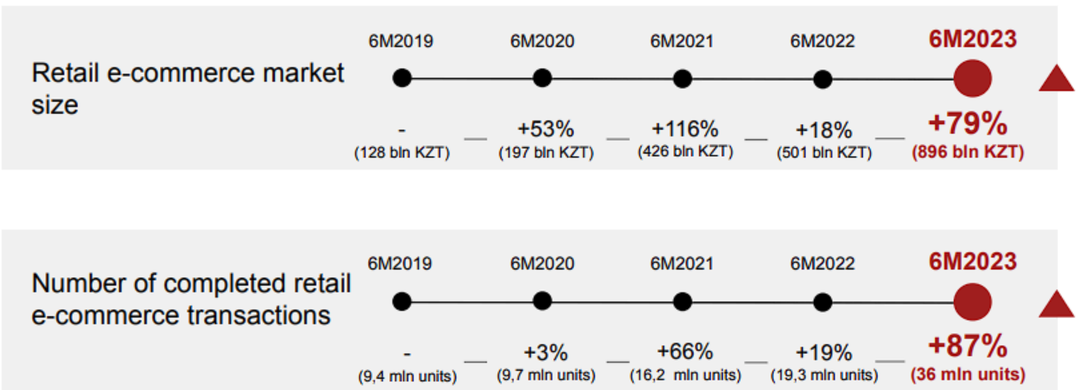

Figure 2. Gathered and Analysed Data of the retail e-commerce market in Kazakhstan [8].

1.1.3 E-commerce among small businesses.

According to information gathered by The New York Times, despite the pandemic having positive effects on e-commerce sales, generally, it affected big businesses. In 2020, 68 percent of the increase in e-commerce sales was accounted for by large retailers [2].

For small businesses, pandemic restrictions caused closures of physical contact between them and customers, which forced many to e-commerce [2].

In 2023, on Main Street, USA, 80 per cent of small and mid businesses were already using online channels, while 16 per cent are also planning to implement them as well. Moreover, on average such businesses generate half of their sales via online channels [9].

In Kazakhstan, 2020, according to Visa-provided research, 63 per cent of micro, small and mid businesses present themselves online via their own sites or social media, or at least use online advertisements [10].

1.1.4 Image presence in the e-commerce retailing sector.

With the invention of photography and mobile phones, images have become an essential part of our lives. In the case of e-sales, we can see the same trend since the start of online sales; images of the products were built on any site (Figure 3).

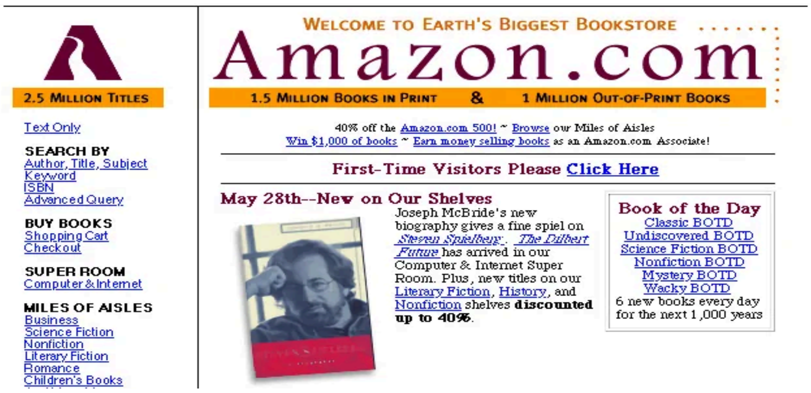

Figure 3. Amazon cropped the homepage image restored by Version Museum (1997) [11].

The essential role of the image of a product is to provide additional information to its description for the consumer, thus facilitating the purchase. Also, the image of a product plays a vital role in the one-click-to-purchase approach used by most retailers, since every additional effort made by the consumer to buy a product decreases its chance to be taken [12].

So all the above mentions business-to-consumer interactions in terms of product pictures. However, there are opposite interactions when the consumer needs to find a specific product in the trader catalogue, without any additional information despite its photo.

Several large traders, such as eBay, Amazon, and others, have already implemented deep-learning technologies for image search tasks. They use different visual search systems adapted to their needs [3], [13].

1.2 Image similarity analysis

Image similarity is a problem of image processing which can be solved using different approaches:

One of them is basic processing such as summing square differences between each pixel's values. Still, such a solution works with, meaningless for human, images, it cannot provide the necessary information in case we are trying to find a similarity between two pictures of dogs or between a dog and a cat. Information lies not in a single pixel but in relations between them [14].

In that case, various deep learning techniques can help. Deep learning networks are multi-layer networks designed to mimic the cerebral cortex, particularly the visual cortex, in our case. There are several techniques that are used in the field of computer vision [15].

1.2.1 Convolutional Neural Network CNN

Convolutional Neural Network (CNN) (Figure 4) is a technique that involves analysing 2d images with a convolution through kernels and assigning weights and biases. During that process, various levels of the neural network find patterns and value combinations that result in different types of images (classes), optimise weights and biases to find them and predict classes more precisely. As a result, through differentiation, it is able to get a class (output) of an image [16].

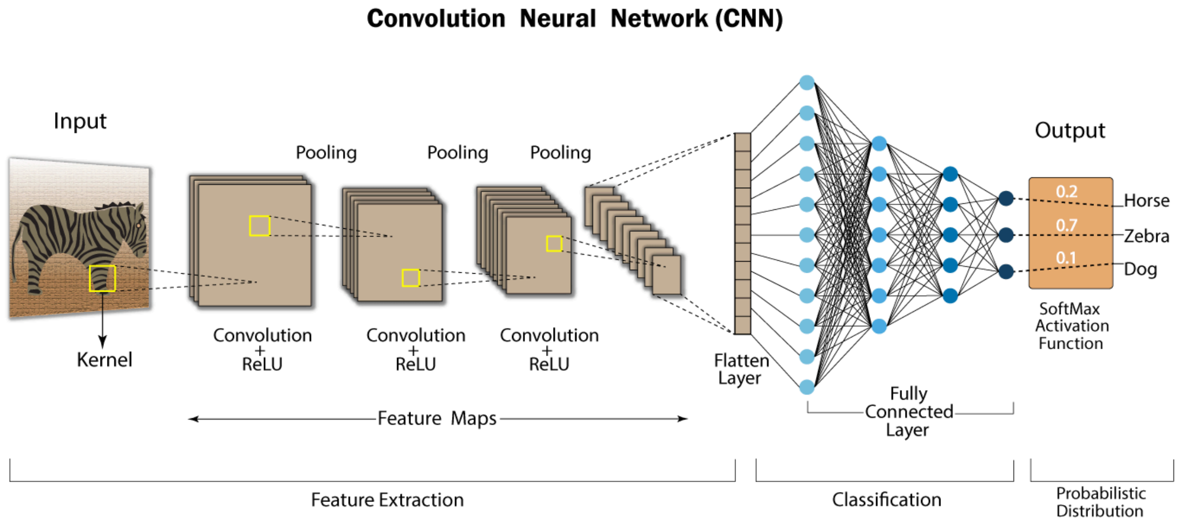

Figure 4. CNN process's example [17].

1.2.2 General Workflow of Neural Network Training.

General workflow of training CNN on example Keras [18]:

1. Data Collection and Preparation.

Gathering Dataset. Find or collect a dataset that fits the testing conditions in terms of both size and content. If we want to train CNN to recognise cats and dogs we need to find a dataset with images labelled as cats and dogs with the same amount for both categories.

Augmentation implementation is the process of applying different, random modifications to images to increase the diversity of available data artificially.

Normalisation. Process of normalising all data in a dataset to a certain standard. Usually, it involves resizing/cropping images to standard dimensions and rescaling pixel values to a common scale.

2. Building a model.

There are two main options: choose a suitable base architecture or design a custom model. Popular CNN architectures are AlexNet, VGGNet, ResNet, DenseNet, and others [19].

After that, the model compilation with Loss Function, Optimizer, and defining Metrics happens. The loss function establishes the quantity of error between the predictions of the model and the actual values [20].

An optimiser is an algorithm used to tweak the model's parameters in pursuit of minimisation of the loss function [21].

Metrics is a method to evaluate the performance of the model. Famous metrics are Accuracy, Precision, Confusion Matrix, Mean Squared Error (MSE), Mean Absolute Error (MAE) and others [22].

3. Training the model.

The dataset must be divided into training and validation parts for the training step. Further division of parts into smaller patches - batches happens. During one batch, part of the data is fed to the model, and loss is calculated, backward calculation of the gradients of learning weights and then the optimiser tweaks learning weights based on gradients. Gradients then zero out, and a new batch is fed for a new cycle. When the whole training dataset is fed to the model, it's called an epoch [23]. Depending on the evaluation step, the number of epochs is determined (Figure 5).

```
n_epochs = 50    # number of epochs to run
batch_size = 10  # size of each batch
batches_per_epoch = len(Xtrain) // batch_size

for epoch in range(n_epochs):
    for i in range(batches_per_epoch):
        start = i * batch_size
        # take a batch
        Xbatch = Xtrain[start:start+batch_size]
        ybatch = ytrain[start:start+batch_size]
        # forward pass
        y_pred = model(Xbatch)
        loss = loss_fn(y_pred, ybatch)
        # backward pass
        optimizer.zero_grad()
        loss.backward()
        # update weights
        optimizer.step()
```

Figure 5. Training loop [24]

4. Evaluation.

After each epoch validation part of the dataset is fed to the model, and metrics is used to determine the quality of its performance [18].

1.2.2.1 Dataset

A dataset is a fundamental structured data source for Neural Network (NN) processing. Mainly, an image dataset is a collection of digital images processed for testing, training, and evaluating the performance of NNs [25].

1.2.2.2 Dataset Processing

Typical steps of processing a dataset before its usage are implementing augmentations and the dataset's normalisation.

Data augmentation is a technique used to increase the training capabilities of a dataset without gathering new data [26]

Standard options available for dataset augmentations are Rotation (Figure 6), Width/Height Shifting, Brightness, Shear Intensity, Zoom, Channel Shift, and Horizontal/Vertical Flip (Figure 7) [27].

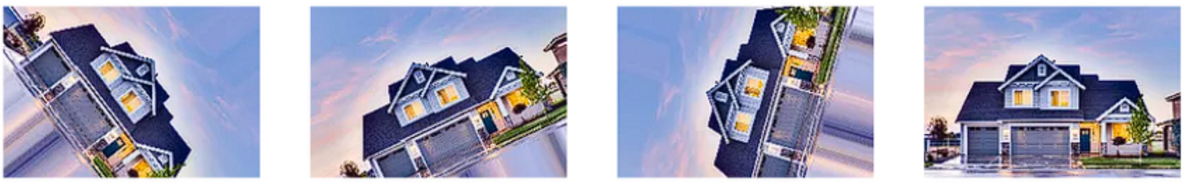

Figure 6. Image augmentation (Rotation) [27].

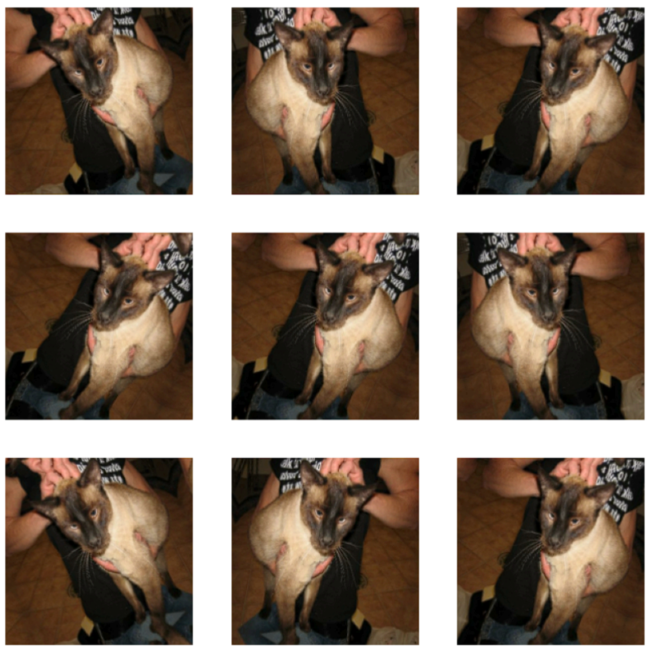

Figure 7. Result of 9 consecutive augmentations (Flip and Rotation) [18]

Such an increase in available data increases the robustness of the model, making it robust to the variability of inputs, like changes in light conditions, camera position, etc.

Augmentations also prevent models from overfitting, when the model "remembers" all training data, and does not perform well with new inputs. Usually, it is observable via differences between training and validation accuracy. Training accuracy becomes ideal, and validation accuracy is on the level of random guesses or even worse [28].

The second phase of processing the dataset before its usage is normalisation. In broader terms, normalisation is reorganising any data to remove any unstructured or redundant data and achieve a standardised data format in the dataset [29].

In terms of CNN normalisation, it usually means reducing RGB channels from [0:255] (integer) to [0:1] (float). As well as cropping/resizing images to standard dimensions [18].

1.2.2.3 Model and its parts

As can be seen in Figure 4, CNN consists of several types of layers, stacked after each other. Depending on the chosen architecture, the number of layers, and their detail differs.

Typical CNN consists of Convolutional, Activation, Pooling, Fully Connected, Dropout, and Normalisation layers [16], [17].

Convolutional layers extract features, via the convolution process, and kernel (mask/filter) is applied through the image, producing a dot product between the filter and input parts (Figure 8) [30].

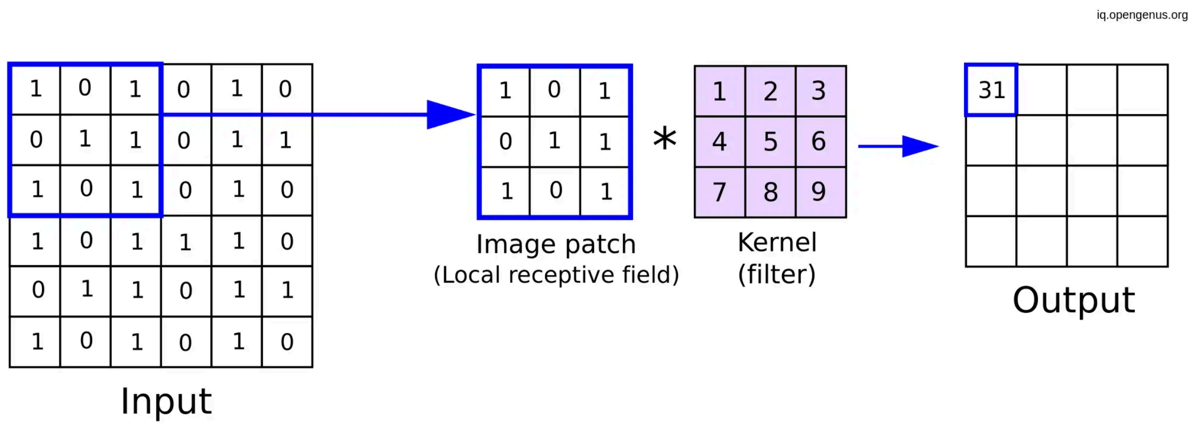

Figure 8. Convolution [30].

Activation layers consist of activation functions, which introduce non-linearity into the model, allowing it to learn complex patterns—common activation functions ReLU, Sigmoid, Tanh (Figure 9).

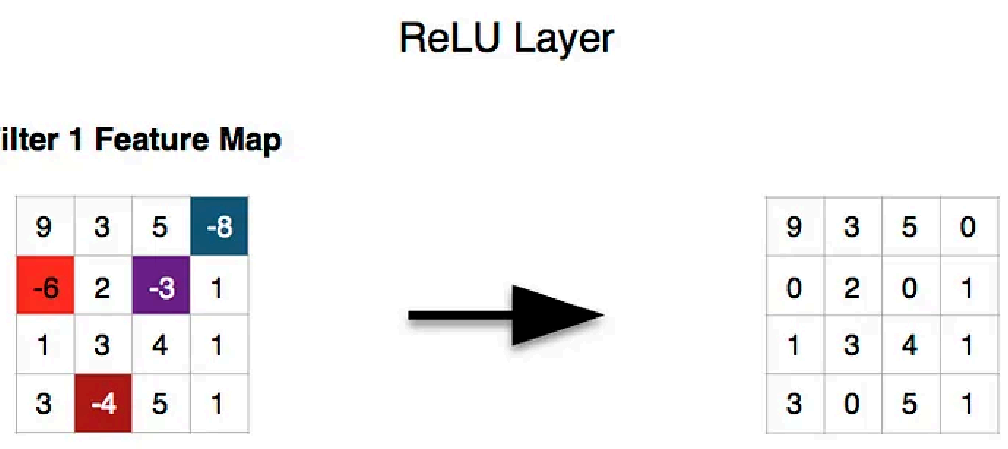

Figure 9. ReLU [31].

Pooling layers have two main types, max pooling and average pooling, their purpose is to reduce the spatial size of the feature maps (Figure 10).

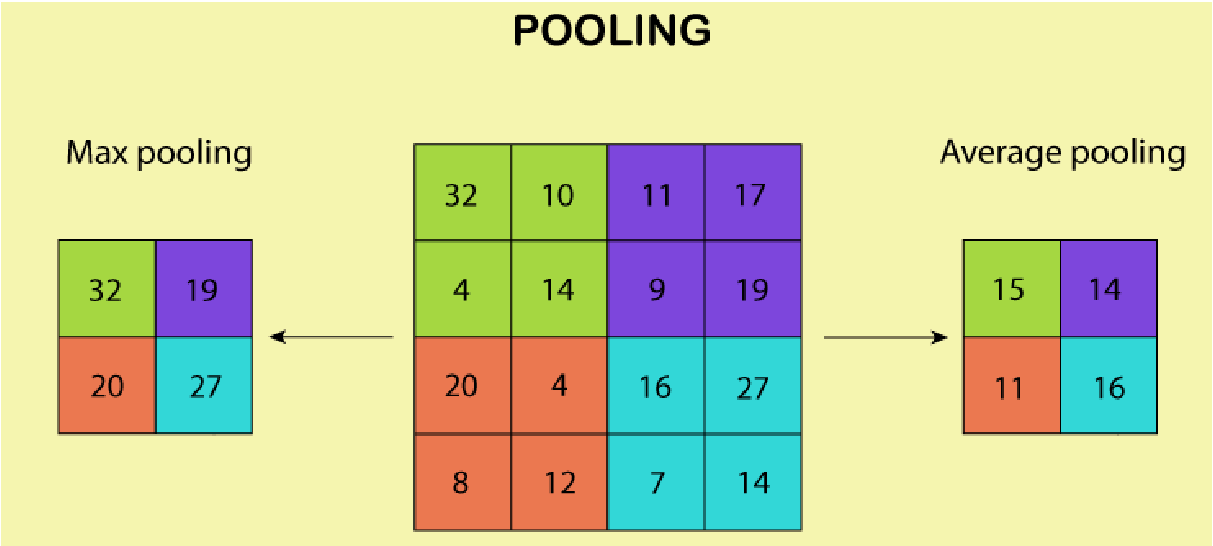

Figure 10 Both Max and Average pooling [32].

Fully Connected layers perform high-level reasoning on the features extracted by convolutional and pooling layers. These layers are similar to those found in traditional neural networks. Using the weight matrix, the neuron applies a non-linear transformation to the input vector, resulting in an output vector (Figure 11).

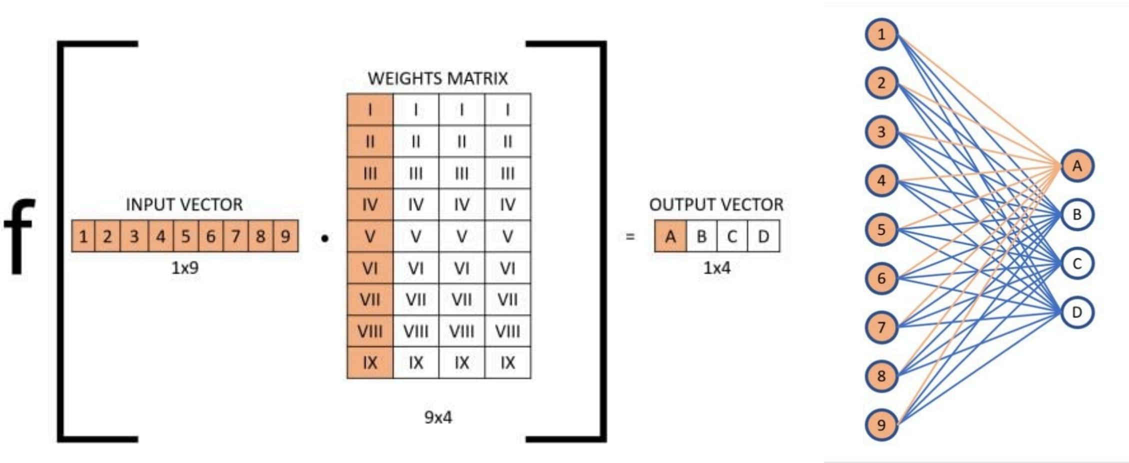

Figure 11 Fully Connected layer [33].

Dropout layers set random input parts to zero, preventing the model from overfitting [34].

Normalisation layers change inputs toward distribution with mean 0 and deviation 1. Stabilise training and speed up the learning process [35].

2 THE AIMS OF THE THESIS

The main goal of the thesis is to evaluate the possibility of implementing neural networks in small businesses for Business-to-consumers interaction. Specifically in the form of an assistant searching for a desired, consumer, product among the business' assortment.

The critical points in the thesis are:

- Ensure validity of the thesis results for small businesses.

It is planned to be achieved by using open-source tools for testing setup, and thus, it is achievable to create and implement in small businesses with limited resources.

Another way to ensure the validity of the result is using comparable data, which is expected to be held by small businesses. So, datasets will be limited in size and content.

- Testing possible approaches to implementing the NNs in B2C types of interactions.

CNN will be tested as a possible solution in assistance in search among business products.

- Evaluation of possible application of CNN for consumer search of products among business' assortment.

CNN will be trained based on a suitable dataset (which could represent data fit for small retail businesses), and the final model will be evaluated via chosen metrics.

3 EXPERIMENTAL PART

3.1 Materials and Methods

3.1.1 Computational environment

All experiments and setup were done on a Windows Subsystem Linux (WSL) Ubuntu virtual machine based on Home Windows 11.

The hardware used for this thesis is a laptop produced by Lenovo, IdeaPad Gaming series, which could be considered a budget class of gaming laptops. In other words, it is available and convenient for small businesses, accounting for its upgradeability and decent specifications [36]. The one used for experiments in this thesis has the following specifications:

AMD Ryzen 7 5800H (3.20 GHz CPU)
NVIDIA GeForce RTX 3050 Laptop GPU

Samsung MZALQ512HBLU-00BL2 (512 GB SSD)

With user upgrades:

Adata Legend 750 (1 TB SSD)

32 GB RAM

For composing the code and conducting experiments, Jupyter Notebook was used. Jupyter Notebook, part of Project Jupyter, is an application for authoring computational notebooks. The computational notebook under Jupyter has a '.ipynb' extension. It is a document that could include computer code, plain texts and visual elements like graphs and figures [37].

Mainly during the coding PyTorch library was used. It is one of the commonly used machine learning libraries [38].

3.1.2 Data and its preparation

The Dataset for the thesis was chosen from Kaggle.com. A Kaggle is a community repository with open access to models, code and, in this thesis case, data [39]. The original dataset was divided into two main groups and four sub-groups. There are 2906 images sorted into four subcategories.

The content and categories of a dataset represent a possible assortment of e-commerce retailing small businesses, with images of clothing divided further into apparel and footwear, which were further divided by men and women.

For experiments, the main categories were eliminated, leaving four subcategories as classes of images, their names: Men, Women, Boys, and Girls (0,1,2,3 respectively (Figure 12)) (Names of sub-categories remained as in source [39]).

The dataset was converted: each image was resized to 224x224 pixels and transformed into PyTorch tensors, and data was converted from [0,255] (integers) into the range [0,1] (float).

Then, the whole dataset was split into training and validation sets, with content distribution of 80 and 20 per cent, respectively.

Horizontal Flip, Rotation and Color Jitter augmentations were used for the experiment. Horizontal flip randomly rotated image around the y-axis. Random rotations up to plus or minus ten degrees were implemented. Color Jitter refers to augmentation, which changes an image's brightness, contrast, saturation, and hue. Brightness randomly changed to power of 0.8-1.2, contrast and saturation by the same factor, and hue adjustment -0.1 or +0.1 (plus or minus 36 degrees shift in hue).

Figure 12. Ten random images from the dataset and their classes.

3.1.3 CNN and its parts

For the experiment pre-trained ResNet-50 (pre-trained on ImageNet dataset), the CNN model was chosen.

All pre-trained layers in the model were frozen. The final layer was replaced with a new fully connected layer with four output features, representing four classes in this thesis' dataset.

Cross Entropy Loss Function was used for the training (Figure 13). Cross entropy loss is used as a common loss function for multiple-class classification problems. Cross Entropy Loss function measures the difference between predicted and result output [40].

$$L\left(y, \hat{y}\right) = -\frac{1}{N}\sum_{i}^{N}\sum_{j}^{C} y_{ij} log\left(\hat{y}_{ij}\right)$$

Figure 13. Categorical Cross Entropy Loss function. N - number of images in a batch, C - Number of classes. yij - predicted result, y^ij - output of the model.

For this thesis, an Adaptive Moment Estimation (Adam) optimiser was chosen. It combines two algorithms, Momentum and Root Mean Square Propagation (RMSP), which allows Adam to pass local minimums efficiently and not overjump global mine in the standard Adam command from PyTorch is the learning rate for the experiments set to 0.0001.

3.1.4 Training

For training, batch sizes were set to 64 images, and epochs were set to 34.

During the training, metric as accuracy was used. After each training batch, validation was used to monitor model training and detect overfitting/underfitting during it.

Both loss and accuracy were recorded during each epoch, and the result was saved as a plot (Figure 14).

As an additional metric, after the training, the confusion matrix was used with the validation dataset as control data (Figure 15).

3.2 RESULTS & DISCUSSION

The Resulted Metrics are saved, and the code is available on Git Hub (Appendix).

During Training, training and validation loss and accuracy were recorded after each epoch. As it can be seen from plots (Figure 14), both loss and accuracy improve with each epoch at a comparable rate, which indicates proper training without under or overfitting.

For Loss, at a random rate with four classes task, the value expected is log(4) or 1.386, the same as the first results after the early epochs of the model. For the final model, both accuracy and loss (near 0.6) results (70-80%) could be considered reasonable.

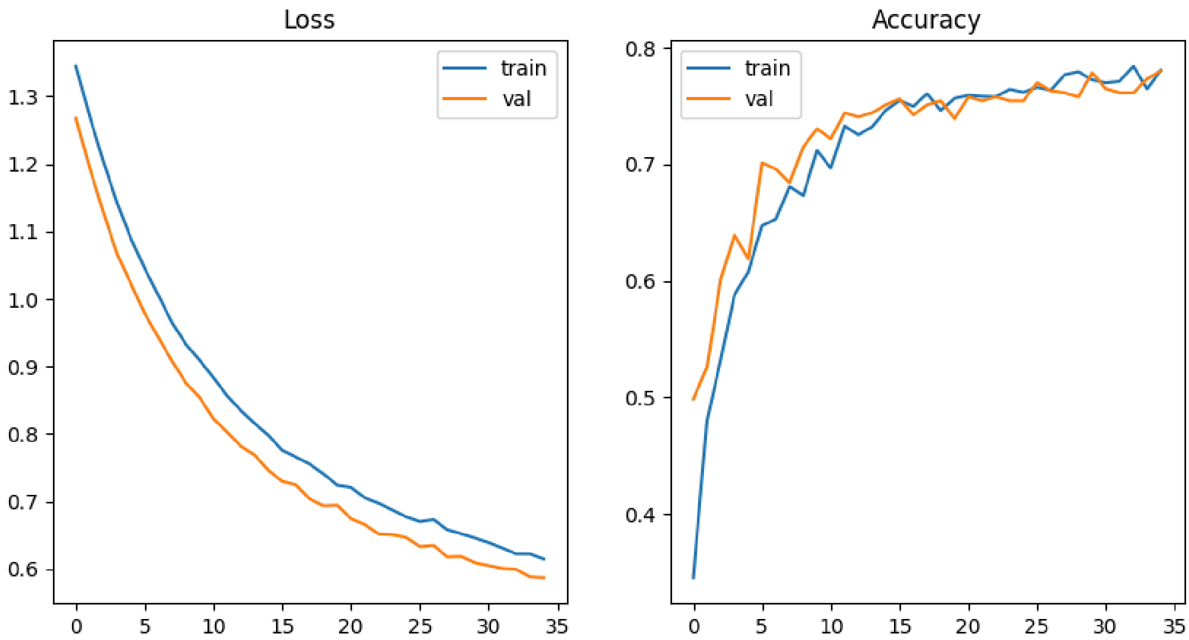

Figure 14. Records of loss and accuracy were obtained during the training of the model.

After completion of the training, the Trained model was fed by the validation dataset, and the results were outlined in the confusion matrix (Figure 15). As expected, the model very rarely confuses large groups (apparel and footwear), and usually makes errors in determining subgroups (Men/Women, Boys/Girls).

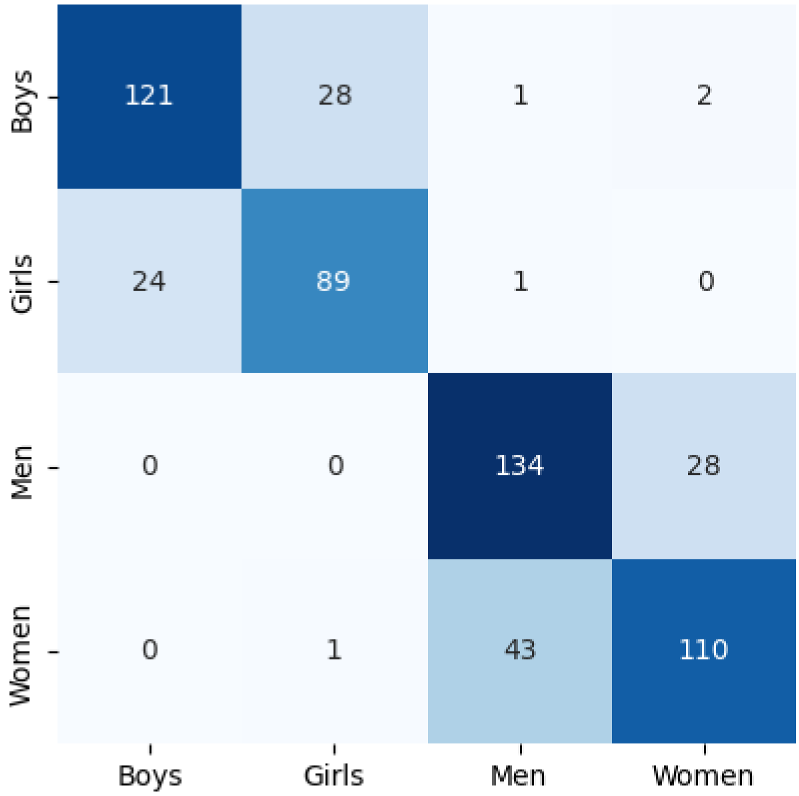

Figure 15. Confusion matrix of validation dataset on the trained model.

In future studies, it is recommended to test other architectures with different parameters.

SUMMARY

In this thesis, the possibility of applying NN in B2C interactions was researched.

CNN was trained on a dataset representing small businesses' theoretically available data, and on achievable by small/micro businesses, hardware.

The fine-tuned ResNet-50 model achieved a 70-80% validation accuracy, with a validation loss of approximately 0.6. These results suggest that NNs can be effectively employed for image-based product searches even with limited data and computational resources. The findings provide a strong foundation for the implementation of neural network-based solutions in small business e-commerce platforms

REFERENCES

https://atameken.kz/ru/news/38948-issledovanie--kazahstanskih-predprinimatelej-dostatochno-aktivno-ispol-zuyut-vozmozhnosti-onlajn-torgovli

[1] ‘35 E-Commerce Statistics of 2024 – Forbes Advisor’. Accessed: May 17, 2024. [Online]. Available: https://www.forbes.com/advisor/business/ecommerce-statistics/

[2] A. Haimerl, ‘When You’re a Small Business, E-Commerce Is Tougher Than It Looks’, *The New York Times*, Mar. 07, 2021. Accessed: May 20, 2024. [Online]. Available: https://www.nytimes.com/2021/03/07/business/small-business-e-commerce.html

[3] M. Du *et al.*, ‘Amazon Shop the Look: A visual search system for fashion and home’, Amazon Science. Accessed: May 17, 2024. [Online]. Available: https://www.amazon.science/publications/amazon-shop-the-look-a-visual-search-system-for-fashion-and-home

[4] V. F. dos Santos, L. R. Sabino, G. M. Morais, and C. A. Gonçalves, ‘E-commerce: a short history follow-up on possible trends’, *E-commerce: um breve histórico de acompanhamento de possíveis tendências*, 2017, doi: 10.5430/ijba.v8n7p130.

[5] J. A. Senn, ‘Electronic Commerce Beyond the “dot com” Boom’, *Natl. Tax J.*, vol. 53, no. 3.1, pp. 373–383, Sep. 2000, doi: 10.17310/ntj.2000.3.04.

[6] K. Jones, ‘The Pandemic Economy: What are Shoppers Buying Online During COVID-19?’, Visual Capitalist. Accessed: May 17, 2024. [Online]. Available: https://www.visualcapitalist.com/shoppers-buying-online-ecommerce-covid-19/

[7] G. Team, ‘Future of E-commerce: Dominance Over Retail by 2024 - Generix Group’, Generix. Accessed: May 17, 2024. [Online]. Available: https://www.generixgroup.com/en/blog/2024-e-commerce-will-overtake-physical-retail

[8] ‘e-commerce-6M2023-eng.pdf’. Accessed: May 22, 2024. [Online]. Available: https://www.pwc.com/kz/en/publications/e-commerce/pdf/e-commerce-6%D0%9C2023-eng.pdf

[9] M. Brohan, ‘Small business bets big on ecommerce’, Digital Commerce 360. Accessed: May 21, 2024. [Online]. Available: https://www.digitalcommerce360.com/2023/12/21/small-business-bets-big-on-ecommerce/

[10] ‘Исследование: 63% казахстанских предпринимателей достаточно активно используют возможности онлайн-торговли’, НПП РК. Accessed: May 21, 2024. [Online]. Available: https://atameken.kz/ru/news/38948-issledovanie--kazahstanskih-predprinimatelej-dostatochno-aktivno-ispol-zuyut-vozmozhnosti-onlajn-torgovli

[11] ‘29 Years of Amazon.com Website Design History - 34 Images’, Version Museum. Accessed: May 17, 2024. [Online]. Available: https://www.versionmuseum.com/history-of/amazon-website

[12] G. L. Lohse and P. Spiller, ‘Electronic shopping’, *Commun. ACM*, vol. 41, no. 7, pp. 81–87, Jul. 1998, doi: 10.1145/278476.278491.

[13] F. Yang *et al.*, ‘Visual Search at eBay’, in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’17. New York, NY, USA: Association for Computing Machinery, Aug. 2017, pp. 2101–2110. doi: 10.1145/3097983.3098162.

[14] M. B. Hisham, S. N. Yaakob, R. A. A. Raof, A. B. A. Nazren, and N. M. Wafi, ‘Template Matching using Sum of Squared Difference and Normalized Cross Correlation’, in *2015 IEEE Student Conference on Research and Development (SCOReD)*, Dec. 2015, pp. 100–104. doi: 10.1109/SCORED.2015.7449303.

[15] T. A. Team, ‘Image Similarity With Deep Learning Explained – Towards AI’. Accessed: May 21, 2024. [Online]. Available:

https://towardsai.net/p/deep-learning/image-similarity-with-deep-learning-explained, https://towardsai.net/p/deep-learning/image-similarity-with-deep-learning-explained
[16] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, 'Convolutional Neural Network (CNN) for Image Detection and Recognition', in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, Dec. 2018, pp. 278–282. doi: 10.1109/ICSCCC.2018.8703316.
[17] D. Kalita, 'Basics of CNN in Deep Learning', Analytics Vidhya. Accessed: May 17, 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2022/03/basics-of-cnn-in-deep-learning/
[18] K. Team, 'Keras documentation: Image classification from scratch'. Accessed: May 21, 2024. [Online]. Available: https://keras.io/examples/vision/image_classification_from_scratch/
[19] 'Convolutional Neural Networks: Architectures, Types & Examples'. Accessed: May 25, 2024. [Online]. Available: https://www.v7labs.com/blog/convolutional-neural-networks-guide, https://www.v7labs.com/blog/convolutional-neural-networks-guide
[20] 'Loss Functions in Machine Learning Explained'. Accessed: May 25, 2024. [Online]. Available: https://www.datacamp.com/tutorial/loss-function-in-machine-learning
[21] '10 famous Machine Learning Optimizers', DEV Community. Accessed: May 25, 2024. [Online]. Available: https://dev.to/amananandrai/10-famous-machine-learning-optimizers-1e22
[22] A. Bajaj, 'Performance Metrics in Machine Learning [Complete Guide]', neptune.ai. Accessed: May 25, 2024. [Online]. Available: https://neptune.ai/blog/performance-metrics-in-machine-learning-complete-guide
[23] 'Training with PyTorch — PyTorch Tutorials 2.3.0+cu121 documentation'. Accessed: May 26, 2024. [Online]. Available: https://pytorch.org/tutorials/beginner/introyt/trainingyt.html
[24] A. Tam, 'Creating a Training Loop for PyTorch Models', MachineLearningMastery.com. Accessed: May 26, 2024. [Online]. Available: https://machinelearningmastery.com/creating-a-training-loop-for-pytorch-models/
[25] '34 Amazing Image Datasets for Your Next CV Project', Datagen. Accessed: May 21, 2024. [Online]. Available: https://datagen.tech/guides/image-datasets/image-datasets/
[26] 'Data augmentation | TensorFlow Core', TensorFlow. Accessed: May 24, 2024. [Online]. Available: https://www.tensorflow.org/tutorials/images/data_augmentation
[27] S. Sarin, 'Exploring Image Data Augmentation with Keras and Tensorflow', Medium. Accessed: May 24, 2024. [Online]. Available: https://towardsdatascience.com/exploring-image-data-augmentation-with-keras-and-tensorflow-a8162d89b844
[28] 'What is Overfitting? - Overfitting in Machine Learning Explained - AWS', Amazon Web Services, Inc. Accessed: May 24, 2024. [Online]. Available: https://aws.amazon.com/what-is/overfitting/
[29] 'Data & Database Normalization Explained: How To Normalize Data', Splunk. Accessed: May 26, 2024. [Online]. Available: https://www.splunk.com/en_us/blog/learn/data-normalization.html
[30] 'Convolutional Neural Networks: 1998-2023 Overview | SuperAnnotate'. Accessed: May 26, 2024. [Online]. Available: https://www.superannotate.com/blog/guide-to-convolutional-neural-networks
[31] H. LaBrie, 'Understanding Autoencoders — Part III: The ReLu Activation Function', Medium. Accessed: May 26, 2024. [Online]. Available: https://blog.gopenai.com/understanding-autoencoders-part-iii-the-relu-activation-function-52b0647c6034

[32] Swapna., ‘Convolutional Neural Network | Deep Learning’, Developers Breach. Accessed: May 26, 2024. [Online]. Available: https://developersbreach.com/convolution-neural-network-deep-learning/
[33] ‘Fully Connected Layer vs Convolutional Layer: Explained | Built In’. Accessed: May 26, 2024. [Online]. Available: https://builtin.com/machine-learning/fully-connected-layer
[34] ‘tf.keras.layers.Dropout | TensorFlow v2.16.1’, TensorFlow. Accessed: May 26, 2024. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Dropout
[35] K. Team, ‘Keras documentation: Normalization layer’. Accessed: May 26, 2024. [Online]. Available: https://keras.io/api/layers/preprocessing_layers/numerical/normalization/
[36] ‘Lenovo IdeaPad Gaming 3 (2021) Review’, RTINGS.com. Accessed: May 26, 2024. [Online]. Available: https://www.rtings.com/laptop/reviews/lenovo/ideapad-gaming-3-2021
[37] ‘Jupyter Notebook Documentation — Jupyter Notebook 7.2.0 documentation’. Accessed: May 26, 2024. [Online]. Available: https://jupyter-notebook.readthedocs.io/en/latest/
[38] ‘PyTorch Foundation’, PyTorch. Accessed: May 26, 2024. [Online]. Available: https://pytorch.org/foundation
[39] ‘E-commerce Product Images’. Accessed: May 26, 2024. [Online]. Available: https://www.kaggle.com/datasets/vikashrajluhaniwal/fashion-images
[40] ‘Cross Entropy Loss Function in Machine Learning - Shiksha Online’. Accessed: May 27, 2024. [Online]. Available: https://www.shiksha.com/online-courses/articles/cross-entropy-loss-function/

Appendix

Model and Jupyter Notebook

The code os accessible in the repository: https://huggingface.co/FriendFromHell/CNNThesis

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

I, Adilet Dossymbekov,

(*author's name*)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Image similarity usage in order to find similar items in E-commerce dataset,

(*title of thesis*)

supervised by Aleksandr Makarov.

(*supervisor's name*)

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

- Adilet Dossymbekov

26/05/2024