Skolkovo Institute of Science and Technology

MSc Data Science, 1st year, Machine Learning

Project Proposal

Pattern mining from text with RNNs

Chesakov Daniil

Glazov Vsevolod

Kovalev Evgeny

Matyushin Leonid

MOSCOW

2019

Abstract

Recurrent neural networks are commonly used in various natural language processing tasks. Their application to sequence modelling problems is an effective way to deal with them, however, there is a lack of clear understanding how to interpret the obtained results, and here pattern mining from texts may help. In this work, we studied activations of the hidden layer neurons of the RNN and LSTM fitted in an unsupervised setting on the parsed item descriptions data from the Etsy website. Finally, we proposed methods for pattern mining in texts in order to identify some entities they contain such as websites or physical metrics of an object.

1 Introduction

In the last few years, deep neural networks have gained great popularity in many tasks of artificial cognition, such as image recognition or text classification. These models all share a common property: they use an implicit representation of the original features without using pre-specified functions studied by a user. These implicit representations have proven to be very efficient for many tasks. However, the black-box nature of these representations make it difficult to interpret the models. Despite the fact that such architectures allow to create efficient systems, understanding of the internal structure is limited and resulting models are poorly interpretable. Visualization and understanding of the hidden layer behavior in RNN can give an impulse for developing new effective tools in task of text analysis.

2 Related work

Our project is mainly based on the paper [1] where authors studied the activation of the hidden layer neurons for character-level multi-layers RNN, LSTM and GRU models. Their work demonstrates that selected cells can detect clear events such as open parentheses, quotes and inline comments in a code.

Also we studied the work [2] which mainly was aimed on developing the GUI visual analyzer for the single- and multi-layers LSTM cell. Authors used the same ideas as in the work [1], but here they used the word-level model in comparison to the character-level model used in [1].

3 Dataset collecting

We decided to collect item descriptions from the Etsy e-commerce website (etsy.com), because it has a good API and most of descriptions are more than 1000 letters. Our dataset contains 10000 unique item descriptions, such as one at the picture.

```
id_ = '567419184'
get_description(id_)
```

```
'A rattle with peas is a traditional rattle consisting of a ball (3 cm in diameter) filled with peas on a beech stick
having three little linden balls. When the baby shakes the rattle, it makes a gentle clacking sound. A unique gift fo
r new parents and their new little ones. Sanded perfectly smooth and covered with linseed oil.\n\nMATERIALS: beech, l
inden, peas. \nLENGTH - 5.1 inch (13 cm)\n100% handmade.\n\nOur toys help to develop:\n-\tfine motion and speaking sk
ills;\n-\timagination and creativity;\n-\ttaste for beauty and love for Nature.\n\nAGE: 0+ YEAR\n\nCraft paper bag in
cluded.\n\nThank you for your visit!\nSee other natural wooden toys https://www.etsy.com/shop/tinyfoxhole\n\nWooden t
oys fill our children with live energy and warmth because they are made of a magic natural material - wood. Being han
d-carved and hand-painted the toys have individual peculiarities, they help to inculcate in children a taste for natu
ral beauty. Wooden toys are simple, but through simple shapes they help children to discover their own creativity, us
e their imagination to the full and compose endless number of games. \n\nNOTE: \n• All of our toys are made from orga
nic, natural materials and non-toxic child safe finishes and paints. \n• This item is handmade and may be slightly di
fferent in wood grain and tone from what you see in pictures, you get a unique piece!'
```

4 RNN architecture

To get unsupervisedly trained model for pattern mining we used character-level language modeling trained to predict the next character in the sequence. In this setting, the input to the network was a sequence of fixed size of one-hot-encoded characters. In order to obtain the prediction we used fully connected layer on top of the RNN output and used Softmax as prediction function (see Figure 1). For training we used Cross-entropy loss.

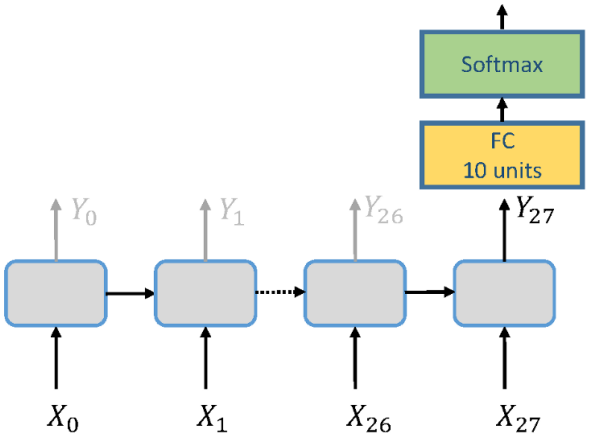

Figure 1: Principal scheme of prediction

RNN

In our experiments we tried a a few different approaches. For the first one we tried the simplest model with RNN cell (shown at Figure 2) in one- and two-layer models. We used RNN cell with 128 neurons inside the hidden layer and train dataset for 30 epochs with batch size 128 and sequence of 1500 characters. We used RMSProp optimizer with step size 0.001.

LSTM

For the second approach we used more complicated RNN cell — Long short-term memory (LSTM) (shown at Figure 3). We also tried one- and two-layer models with this type of cells and same parameters as in previous case.

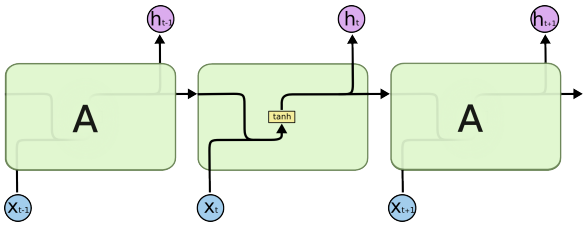

Figure 2: RNN cell

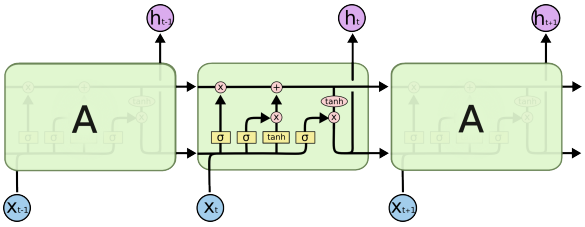

Figure 3: LSTM cell

5 Pattern mining

With our character-level model, we were able to extract activation vector corresponding to each character in the given text. More specifically, given the previous hidden state and the input character, at each step our RNN model generates a new hidden state which we call an activation vector. So, at the end of text processing we obtain a sample of activation vectors $A \in \mathbb{R}^{n \times k}$, where n is a number of characters in text and k is the length of activation vector (which is also a number of hidden neurons in network). By taking some component of activation vectors for all characters in text (some particular column in A), we can look for the patterns in it.

However, these patterns are still not well recognizable. Figure 4 shows an example of series of values for one of activation components for some text. Due to many fluctuations, it is really hard to see some patterns in it. To overcome this issue, we printed texts in a colored manner. More particularly, we normalized activation components to take values from 0 to 1 and then applied colormap to each character which colored values near 0 with dominant blue, near 0.5 with dominant red and near 1 with dominant green.

In the majority of studied cases the application of such procedure for some activation component showed no explicit pattern. Nevertheless, some activation components clearly caught some dependencies. Figure 5 shows an example of how RNN was able to learn first letters of words in the text, and Figure 6 shows an example

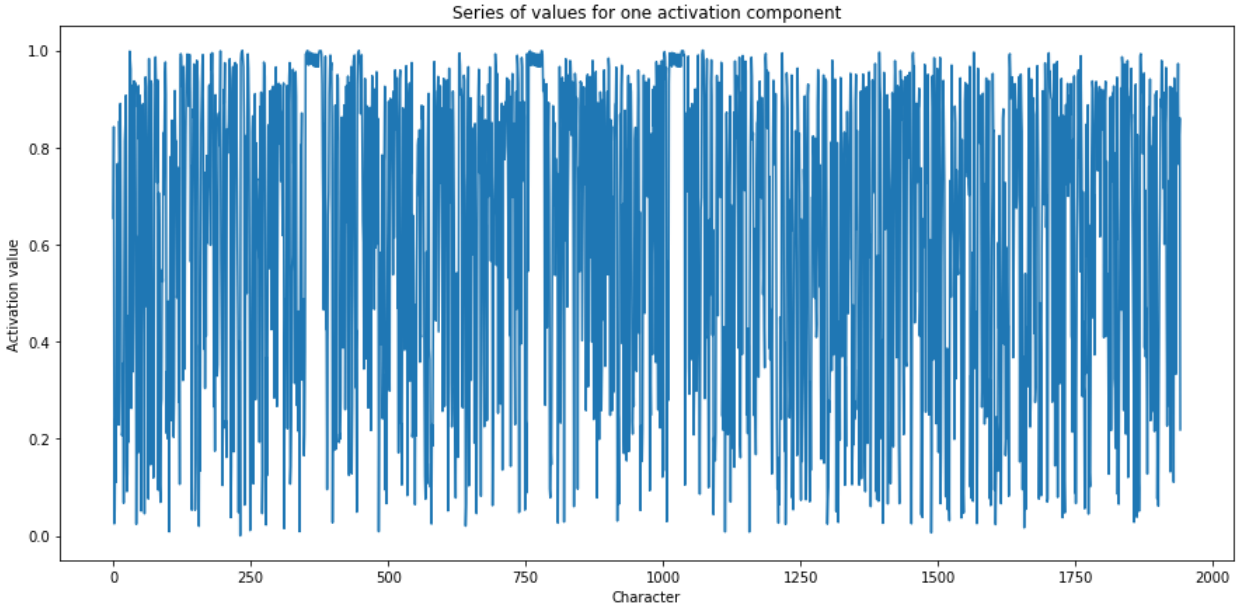

Figure 4: Series of values for one of activation components for an example text

This lovely doll is made with gathered natural materials like sticks and has a terracotta clay face. All her clothes, hair and jewels are repurposed and recycled. This hand made spirit doll is nameless until it finds its way to its owner, which is a woman. Then she can reflect on her inner heart and decide what spirit or name the doll represents. Once that is done the doll can help you find the ability to unlock what is keeping you stuck and to see and perhaps change what no longer benefits you. The hidden message in this doll is love, peace in the new year. Your spirit doll will support you and help you make decisions that are beneficial to you. Generally spirit dolls are something a person makes with their own two hands, however these dolls are made and designed to be a receptacle that is open to their owner once found. This lovely doll has a strong connection to water. Stand 15" tall

Figure 5: First letters of words by RNN activation

This lovely doll is made with gathered natural materials like sticks and has a terracotta clay face. All her clothes, hair and jewels are repurposed and recycled. This hand made spirit doll is nameless until it finds its way to its owner, which is a woman. Then she can reflect on her inner heart and decide what spirit or name the doll represents. Once that is done the doll can help you find the ability to unlock what is keeping you stuck and to see and perhaps change what no longer benefits you. The hidden message in this doll is love, peace in the new year. Your spirit doll will support you and help you make decisions that are beneficial to you. Generally spirit dolls are something a person makes with their own two hands, however these dolls are made and designed to be a receptacle that is open to their owner once found. This lovely doll has a strong connection to water. Stand 15" tall

Figure 6: Sentence beginning by LSTM activation

This listing is for the INSTANT DOWNLOAD KNITTING PATTERN of "The Triangle Scarf" by Bo Peep's Bonnets. This pattern is available for PERSONAL, noncommercial use ONLY. No unauthorized reproduction or distribution, in whole or part of this pattern is permitted. Items created or derived from this pattern are not allowed to be sold in any capacity. The pattern and photographs contained in this document are the property of Bo Peep's Bonnets©. * All patterns written in standard US terms. * Size: Teen/Adult * Any Super Bulky Yarn * Additional photos to help with button sewing and I-cord attachment *There is no shipping charge for this item as it is a PDF file and will be available for instant download after purchased. Once you payment is received, Etsy will automatically prompt you to download the file. You can also go to "my account" then "purchases" for the same download link at any time after payment. If you have any trouble please contact me* !!!!Price is for the PATTERN only, not the finished product!!!! Want to purchase the actual infinity scarf? It's available in lots of beautiful colors here - https://www.etsy.com/listing/198538253/hand-knit-triangle-scarf-neutral-oatmeal?ref=shop_home_active_10 * Facebook: https://www.facebook.com/BoPeepsBonnets * Pinterest: http://www.pinterest.com/BoPeepsBonnets/bo-peep-s-bonnets/ * Instagram: www.instagram.com/bopeepsbonnets * Twiter: www.twitter.com/bopeepsbonnets * Visit Bo Peep's Bonnets official website to learn about her charity for cancer patients: http://bopeepsbonnets.com/

Figure 7: Websites by RNN activation

This listing is for the INSTANT DOWNLOAD KNITTING PATTERN of "The Triangle Scarf" by Bo Peep's Bonnets. This pattern is available for PERSONAL, noncommercial use ONLY. No unauthorized reproduction or distribution, in whole or part of this pattern is permitted. Items created or derived from this pattern are not allowed to be sold in any capacity. The pattern and photographs contained in this document are the property of Bo Peep's Bonnets©. * All patterns written in standard US terms. * Size: Teen/Adult * Any Super Bulky Yarn * Additional photos to help with button sewing and I-cord attachment *There is no shipping charge for this item as it is a PDF file and will be available for instant download after purchased. Once you payment is received, Etsy will automatically prompt you to download the file. You can also go to "my account" then "purchases" for the same download link at any time after payment. If you have any trouble please contact me* !!!!Price is for the PATTERN only, not the finished product!!!! Want to purchase the actual infinity scarf? It's available in lots of beautiful colors here - https://www.etsy.com/listing/198538253/hand-knit-triangle-scarf-neutral-oatmeal?ref=shop_home_active_10 * Facebook: https://www.facebook.com/BoPeepsBonnets * Pinterest: http://www.pinterest.com/BoPeepsBonnets/bo-peep-s-bonnets/ * Instagram: www.instagram.com/bopeepsbonnets * Twiter: www.twitter.com/bopeepsbonnets * Visit Bo Peep's Bonnets official website to learn about her charity for cancer patients: http://bopeepsbonnets.com/

Figure 8: Websites by LSTM activation

100mm X 100mm Dragon design. 3mm think, clear coated with a rubberized coating.

Figure 9: Item sizes by RNN activation

100mm X 100mm Dragon design. 3mm think, clear coated with a rubberized coating.

Figure 10: Item sizes by LSTM activation

of how LSTM learned to identify the beginning of sentences. It is also possible to extract links to websites from texts, as shown on Figures 7 and 8. Another interesting example is shown on Figures 9 and 10 — some activation components were able to identify item sizes among the whole description, which may be useful for some applications and doesn't require regular expressions.

Finally, one may notice that it can be time-consuming to go through all the activation components in order to find some patterns. Therefore we applied dimensionality reduction method PCA to matrix A in order to transform it to some matrix $B \in \mathbb{R}^{n \times d}$ where $d << k$. As a result it is possible to go through much less number of components d (say, $d = 3$ or $d = 5$) and find some pattern with their help. You may find the examples on Figures 11, 12 and 13.

It is also worth noting that two-layer models didn't give more interpretable patterns and generally worked pretty much the same as one-layer models.

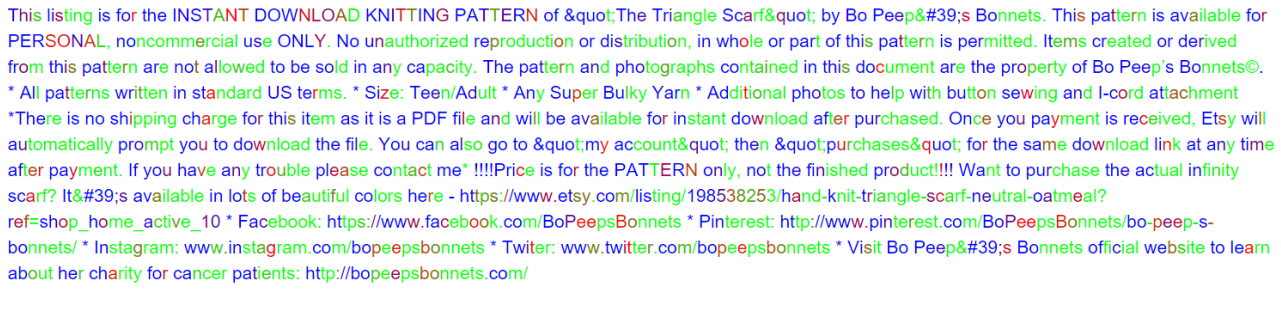

Figure 11: Color patterns by RNN activation after PCA

Figure 12: Color patterns by LSTM cell activation after PCA

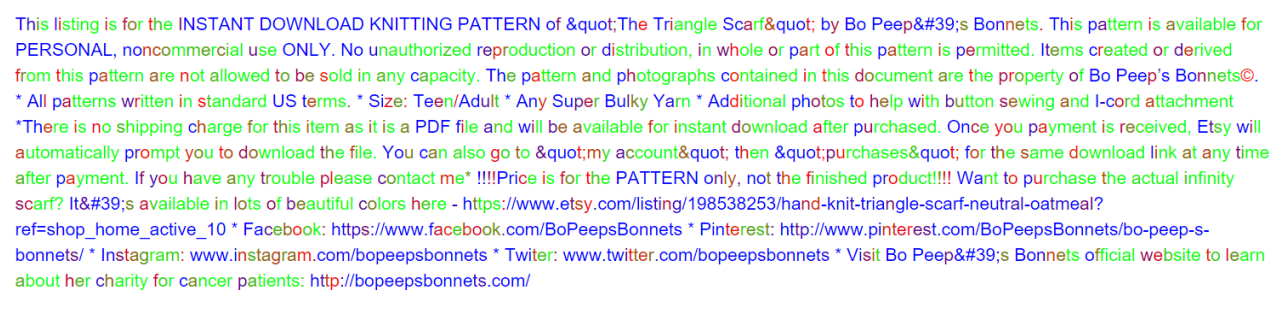

Figure 13: Color patterns by LSTM hidden activation after PCA

6 Binary classification for pattern recognition

In the previous section it was shown that Recurrent Neural Networks actually could learn some information about patterns in the text (in a broad sense). Moreover, as we have seen, information about such patterns contains inside of hidden layers activations. So we can actually train a binary classifier that could detect such patterns.

1. Train a Char-RNN on corpus (next symbol prediction). Cheap step, we do not need any data markup.
2. Perform a data markup of symbols in the corpus (**1** - symbol is inside of a pattern, **0** - symbol is outside of a pattern)
3. $\mathbf{X}$ is a set of activations, computed on symbols from corpus. $\mathbf{y}$ was created on the previous step
4. Train your favorite binary classifier on ($\mathbf{X}$, $\mathbf{y}$)

In our experiment we have considered a dataset that was specified in section 3. We have fitted a 3 layers LSTM on this dataset. After this we put a fixed substring inside of this dataset. We have computed activations of the third layer on our artificial pattern and random non-pattern objects. For example, it is clear that component #19 of the input LSTM gate actually can determine such a pattern, as shown on Figure 14.

We have trained a Support Vector Classifier on the whole input gate vector. We have got a 92.5% accuracy on the test set.

This problem was more or less artificial, since we don't have any resources to perform a massive data markup to solve a real problem. However, as it was shown in the previous sections, Recurrent Neural Networks could

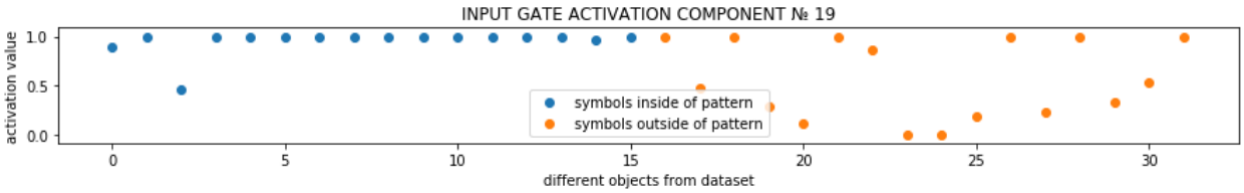

Figure 14: Pattern recognition with LSTM activation

easily determine long relationships between letters in the text. So this approach could actually work in a more general setting.

7 Conclusion

We studied activations of the hidden layer neurons of the RNN and LSTM. We used character-level language models trained on unclear real world data and developed a methodology for mining parts of text that correspond to some frequent patterns such as web cites or physical metrics of an object. Further work may include pattern mining with additional layers, such as convolutional or attention.

8 Team Contribution

- Chesakov Daniil: data parsing and mining
- Glazov Vsevolod: model architecture design
- Kovalev Evgeny: pattern mining with activation components analysis
- Matyushin Leonid: Formal Binary Classification approach to the Pattern Recognition problem

9 GitHub

Here you can find a code to reproduce our experiments:
`https://github.com/blacKitten13/Pattern-Mining-from-text-with-RNNs`

References

[1] A. Karpathy, J. Johnson, and F.-F. Li. Visualizing and understanding recurrent networks. *arXiv preprint arXiv:1506.02078*, 2015.

[2] Hendrik Strobelt, Sebastian Gehrmann, Hanspeter Pfister, and Alexander M. Rush. LSTMVis: A Tool for Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks. *arXiv preprint arxiv:1606.07461*, 2017

[3] M.D.Zeilerand, R.Fergus. Visualizing and understanding convolutional networks. *In European Conference on Computer Vision*, pp. 818–833. Springer, 2014