ERG3020 REPORT

ZHUOYU LI& BOKAI XU& XIYAN LUO

2021.5.5

CONTENTS

ABSTRACT

This project is a demo of social network comment section with a special algorithm on the sorting and the presentation of comments. We aim at find and tell the truth so that we combine Natural Language Processing and Markov Networks to make a demo of social network comment section and we hope that a better social network comment section can be guaranteed.

* *School of Data Science, The Chinese University of Hong Kong, Shenzhen, China*

1 INTRODUCTION

1.1 The Properties of Markov Network and the Origin of Our Idea

Markov Network is an undirected graphical model for representing dependencies between random variables.

A Markov network can be represented by an undirected graph $G = (V, E)$ where the nodes in V represent random variables and edges in E represent dependency relationships.

Let us consider X, a kind of assignment of values to the variables in a Markov Network. We call C the set of maximal cliques in the network and assign a factor ψ_c to each clique $c \in C$. The probability $P(X)$ is,

$$P(X = x) = \frac{1}{Z} \prod_k \phi_k(x_{(k)}) \tag{1}$$

We know that Markov Network can be used to represent a system, where this system is in general form, because the random variables in V can be either numerical or categorical.

If we want to represent a world in the form of Markov Network, we can consider binary case, for example, if Tom lied, we can assign **True** to the random variable **lie(Tom)**. So each random variable in our desired network can only take one of the two binary values **{True, False}**.

The world above is an assignment of values of V.

We are continuously thinking about the way to infer the truth of the event. PageRank algorithm by Jimmy Page gave us the confidence to utilize graph theory to solve the truth of an event. We tried Pagerank Algorithm to infer the truth of an event by assigning the relationship **{Entailment, Contradiction, Independence}** between any pairs of comments on the social network. Then we construct a directed graph of this set of comments. Each comment will post its importance to the comment that it entails, and we can construct a transition matrix, and then calculate A^500 or above, until the matrix converges. Then we can find the stationary point. Then we can rank the comments according to their PageRank value. In this process, we observe the specfic property of graph. However, this algorithm may not work very well in practice, because its time complexity is $O(n!)$. Now we focus on the combination of graph and First Oerder Logic.

We find that undirectional graph may perform better than directional graph. Because the computer cannot really understand what you mean at a extremely high precision, and the sentences generated by users may not entail each other by our definition. So, to use First Order Logic and build an undirectional graph may be better than PageRank algorithm when we analyse the comments by netizens.

1.2 Why We Use The First Order Logic

Atom clauses can be easily understood by programming languages, and it is similary to functions and thus can be easily processed by algorithms. In this project, we aim to convert every natural language sentence into atom clauses of First Order Logic. For example, **Tom accuses Bob of stealing** can be converted into First Order Logic expression **accuseofstealing Bob,Tom**. And this form can be utilized by Markov Network as we mentioned above, because **accuseofstealing**$\mathtt{Bob, Tom}$ can be assigned a value between **{True, False}**. The combination of First Order Logic and Markov Network can utilize the strength of two models.

However, we have to mention that the First Order Logic can be used with Markov Network in two ways: The first role is the representation of events. The second role is to help us determine the possible Markov Network. Why? Because the assignment of values of Markov Network is not known. We have to determine which possible way of assignment is more probable. Then we consider some constrains. These **contrains** are also written in First Order Logic form and can help to contruct Markov **Logic** Network.

1.3 The Nature of Markov Logic Networks

We can view Markov Logic Network (MLN) as the template of Markov Network. MLN is underdetermined and uncertain. We need to determine the variable value. However, because the value is difficult to determine, MLN uses constrains in First Order Logic (FOL) to select those most feasible world. Worlds with more conflicts with stronger constrains has a low probabilty of existence. So MLN meets our demand very well.

1.4 The Components of Markov Logic Networks

Before we get started, we shall first define the components of MLN. MLN consists of four parts represented in FOL:

1. **Knowledge Base** *Some facts about this possible world, written in FOL. For example, "Tom accuses Bill of stealing the money" is represented as "accuseofstealingmoney*$\mathtt{Tom, Bill}$*".*
2. **Function Declaration** *Some functions mentioned in the world, written in FOL. For example, "accuseofstealing*$\mathtt{person, person}$*". This function has two inputs, the first argument is a person and the second argument is also a person.*
3. **Named Entities** *Some entities that belong to some specific categories. For example, "person = {Tom, Jerry, Bob, Nazarbayev}".*

4. **Predicates** Constrains *Some general rules without pointing out any named entities, objects in the constrians are represented as x,y,z, and some underdetermined objects. For example, stealthemoney*$x\hat{}$*accuseofstealing*x, y *then lie*y *means "if x stole the money and y did not accuse x of stealing money y lied."*

2 KNOWLEDGES USED IN THE PROJECT

2.1 Natural Language Processing

Natural language processing (NLP) is a field concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data. The result is a computer capable of "understanding" the contents of documents, including the contextual nuances of the language within them [?].

We have to deal with the natural language records so that we have to choose some NLP algorithms. The most important part is to do text segmentations, transfer the texts into first order logic and atomic sentences to match the requirements of the Markov Logic Network.

2.2 Markov Logic Networks

The markov networks model used in this project comes from the article written by Richardson and Domingos in 2006 [?]. The below statements in this subsection are from the article.

A Markov network (also known as Markov random field) is a model for the joint distribution of a set of variables $X = (X_1, X_2, \cdots, X_n) \in X$. The joint distribution represented by a Markov network is given by

$$P(X = x) = \frac{1}{Z} \prod_k \phi_k(x_{(k)}) \tag{2}$$

where $x_{(k)}$ is the state of the kth clique (i.e., the state of the variables that appear in that clique). Z, known as the *partition function*, is given by $Z = \sum_{x \in X} \prod_k \phi_k(x_{(k)})$.

Definition 1 (Markov logic network)**.** A Markov logic network L is a set of pairs (F_i, w_i), where F_i is a formula in first-order logic and w_i is a real number. Together with a finite set of constants $C = c_1, c_2, \cdots, c_{|C|}$, it defines a Markov network $M_{L,C}$ (Equations 1 and 2) as follows:

1. $M_{L,C}$ *contains one binary node for each possible grounding of each predicate appearing in* L. *The value of the node is* 1 *if the ground atom is true, and* 0 *otherwise.*

2. $M_{L,C}$ *contains one feature for each possible grounding ofeach formula* F_i *in* L. *The value ofthis feature is* 1 *if the ground formula is true, and* 0 *otherwise. The weight ofthe feature is the wi associated with* F_i *in* L.

All the formulas and the constants in the Markov Logic Networks have to meet the 3 assumptions below:

1. **Unique names.** *Different constants refer to different objects.*
2. **Domain closure.** *The only objects in the domain are those representable using the constant and function symbols in* (L, C)
3. **Known functions.** *For each function appearing in* L*, the value of the function applied to every possible tuple of arguments is known, and is a element of* C.

3 PACKAGES USED IN THE PROJECT

3.1 Natural Language Processing

The NLP package used in this project is **AllenNLP** [**?**]. We use the methods of AllenNLP to do text segmentations, transfer the texts into first order logic and atomic sentences so that we can match the requirements from the Markov Logic Network.

3.2 Markov Logic Networks

Pracmln is a toolbox for statistical relational learning and reasoning and as such also includes tools for standard graphical models [**?**]. We use this package to build the markov logic networks and give the results.

3.3 Flask

Flask is a micro web framework written in Python [**?**]. We use **Flask** to show the demo of social network comments section. Here is a figure of the demo website:

On the website, we can find that we can input and post comments in a text area. We even can choose the type of the comments. However, if we choose the wrong type, the NLP model will indentify it and change the type into the correct one. All of the comments will be showed below the text area. Also, the names and functions in the comments will be showed in the left sides. The result of the Markov Logic Networks will be showed below them.

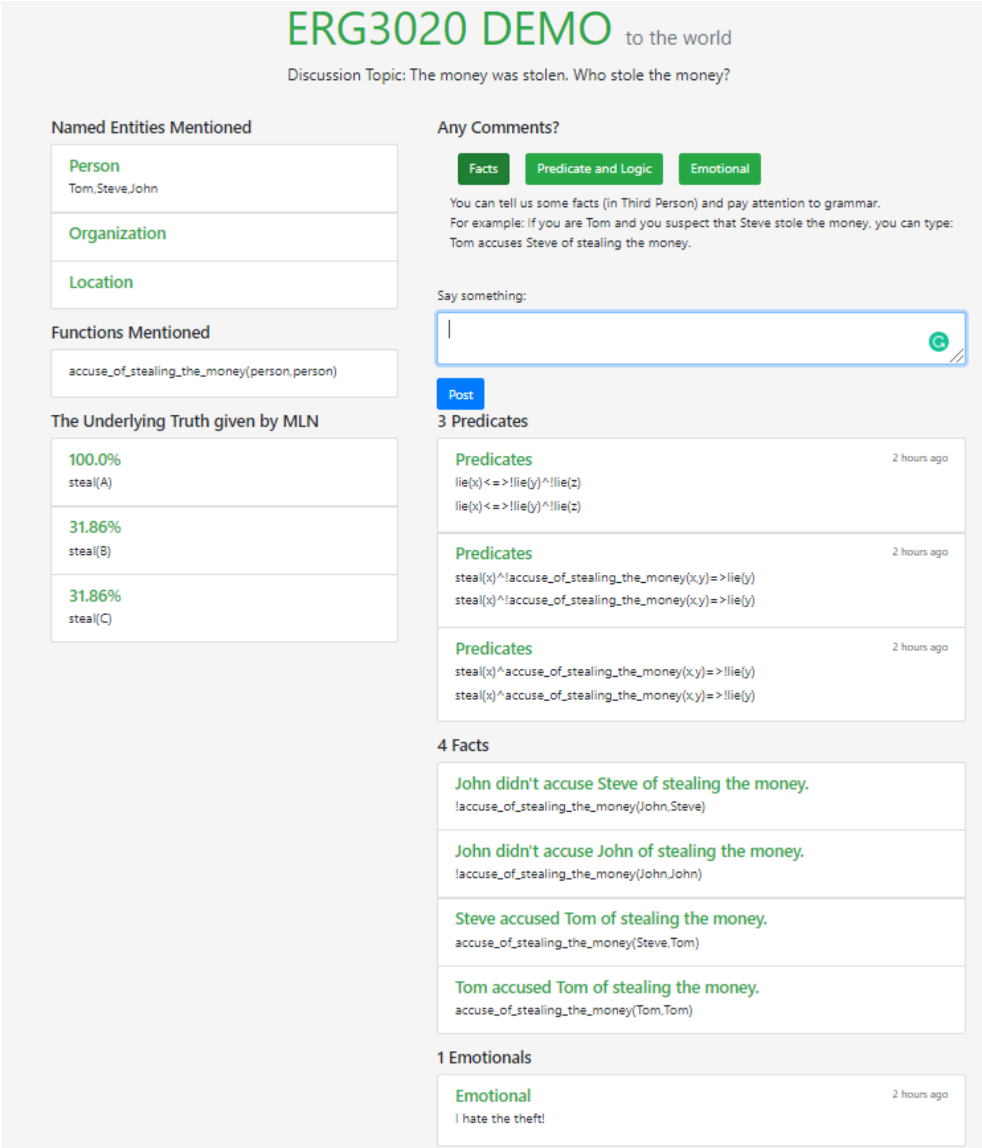


Figure 1: An Screenshot of the website

4 EXAMPLE

Here we show a example with full steps of the usage of the demo.

5 CONCLUSION AND FURTHER DEVELOPMENT

6 INTRODUCTION

A statement requiring citation.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Some mathematics in the text: $\cos\pi = -1$ and α.

7 METHODS

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo.

Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

1. First item in a list
2. Second item in a list
3. Third item in a list

7.1 Paragraphs

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

PARAGRAPH DESCRIPTION Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

DIFFERENT PARAGRAPH DESCRIPTION Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

7.2 Math

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

$$\cos^3 \theta = \frac{1}{4} \cos \theta + \frac{3}{4} \cos 3\theta \tag{3}$$

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Definition 2 (Gauss)**.** To a mathematician it is obvious that $\int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}$.

Theorem 1 (Pythagoras)**.** *The square of the hypotenuse (the side opposite the right angle) is equal to the sum of the squares of the other two sides.*

Proof. We have that $\log(1)^2 = 2\log(1)$. But we also have that $\log(-1)^2 = \log(1) = 0$. Then $2\log(-1) = 0$, from which the proof. $\square$

8 RESULTS AND DISCUSSION

Reference to Figure **??**.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

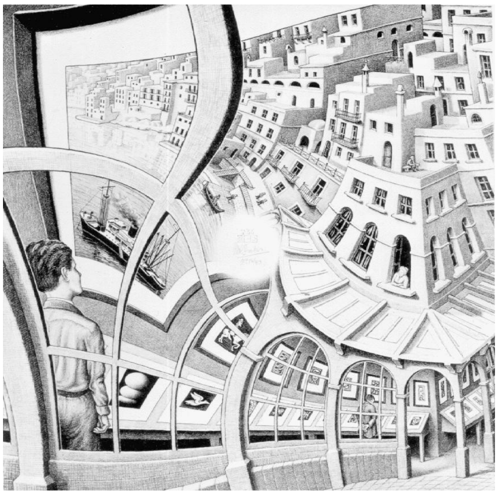


Figure 2: An example of a floating figure (a reproduction from the *Gallery of prints*, M. Escher, from http://www.mcescher.com/).

8.1 Subsection

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

8.1.1 *Subsubsection*

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

WORD Definition

CONCEPT Explanation

IDEA Text

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

- First item in a list
- Second item in a list
- Third item in a list

8.1.2 *Table*

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

Table 1: Table of Grades

Name		
First name	Last Name	Grade
John	Doe	7.5
Richard	Miles	2

Reference to Table **??**.

8.2 Figure Composed of Subfigures

Reference the figure composed of multiple subfigures as Figure **??**. Reference one of the subfigures as Figure **??**.

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

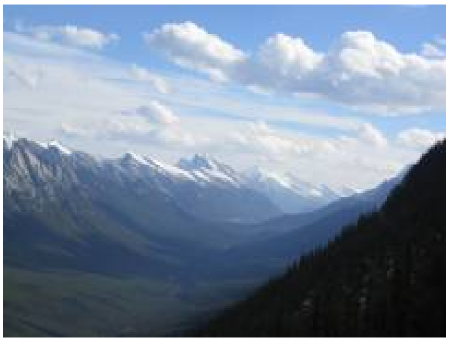
Nulla mattis luctus nulla. Duis commodo velit at leo. Aliquam vulputate magna et leo. Nam vestibulum ullamcorper leo. Vestibulum condimentum rutrum mauris. Donec id mauris. Morbi molestie justo et pede. Vivamus eget turpis sed nisl cursus tempor. Curabitur mollis sapien condimentum nunc. In wisi nisl, malesuada at, dignissim sit amet, lobortis in, odio. Aenean consequat arcu a ante. Pellentesque porta elit sit amet orci. Etiam at turpis nec elit ultricies imperdiet. Nulla facilisi. In hac habitasse platea dictumst. Suspendisse viverra aliquam risus. Nullam pede justo, molestie nonummy, scelerisque eu, facilisis vel, arcu.

Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam congue neque id dolor.

Donec et nisl at wisi luctus bibendum. Nam interdum tellus ac libero. Sed sem justo, laoreet vitae, fringilla at, adipiscing ut, nibh. Maecenas non sem quis tortor eleifend fermentum. Etiam id tortor ac mauris porta vulputate. Integer porta neque vitae massa. Maecenas tempus libero a libero posuere dictum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aenean quis mauris sed elit commodo placerat. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Vivamus rhoncus tincidunt libero. Etiam elementum pretium justo. Vivamus est. Morbi a tellus eget pede tristique commodo. Nulla nisl. Vestibulum sed nisl eu sapien cursus rutrum.

(a) A city market.

(b) Forest landscape.



(c) Mountain landscape.

(d) A tile decoration.

Figure 3: A number of pictures with no common theme.

REFERENCES

[1] Wikipedia contributors. Natural language processing — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Natural_language_processing&oldid=1020063620, 2021. [Online; accessed 5-May-2021].

[2] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine learning*, 62(1-2):107–136, 2006.

[3] Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson F. Liu, Matthew Peters, Michael Schmitz, and Luke S. Zettlemoyer. Allennlp: A deep semantic natural language processing platform. 2017.

[4] Daniel Nyga, Mareike Picklum, Michael Beetz, et al. pracmln – markov logic networks in Python, 2013–. [Online; accessed <date>].

[5] Miguel Grinberg. *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.", 2018.

A A GUIDENCE TO RUNNING CODE

First we have to make sure the working director of the ternimal is the code folder. Then we type the below codes:

```
pip install -r requirements.txt
```

Then we can find that all the requirements are installed. The next steps are to run the flask app.

```
$env:FLASK_APP = "sayhello"
```

And then

```
flask run --host 127.0.0.1 -p 80
```

Therefore, the flask app is run and we can see the website.