

# Implementing of RankNet

Ryan Greenup

February 15, 2021

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Motivation</b>	<b>2</b>
<b>3</b>	<b>Implementation</b>	<b>2</b>
3.1	Blobs . . . . .	2
3.2	Moons . . . . .	2
3.3	Optimisers . . . . .	2
3.4	Batches . . . . .	2
3.5	Wine . . . . .	2
3.6	Rank Wiki Articles . . . . .	2
<b>4</b>	<b>Conclusion</b>	<b>2</b>
<b>5</b>	<b>Further Research</b>	<b>3</b>
<b>6</b>	<b>Text and References</b>	<b>3</b>
<b>7</b>	<b>Fractals</b>	<b>3</b>
	# #+TODO: TODO IN-PROGRESS WAITING DONE	

### 1 Introduction

Ranknet is an approach to *Machine-Learned Ranking* developed by Microsoft as early as 2004 [2], although previous work in this area had already been undertaken as early as the 90s (see generally [4, 3, 4, 5, 11]) these earlier models didn't perform well compared to more modern machine learning techniques [7, §15.5].

*Machine-Learned Ranking* is the process of applying machine learning algorithms to ranking problems, it often referred to as *Learning to Rank*/ [6].

### 2 Motivation

Much data cannot clearly categorised or quantified even if there is a capacity to compare different samples, the motivating example is a collection of documents, it might be immediately clear to the reader which documents are more relevant than others, even if the reader would not be able to quantify a "relevance score" for each document.

By training a model to identify a more relevant document, a ranking can be applied to the data.

An example of this might be identifying documents in a company's interwiki that are relevant for new employees, by training the model to rank whether one document is more relevant than another, ultimately an ordered list of documents most relevant for new employees could be created.

### 3 Implementation

To implement the RankNet approach, a neural network will be trained to identify a score.

Then a sigmoid function to identify a similarity score.

First apply to an obvious data set.

Then use a different optimiser Then allow it to use more data then a different data set Then a non-obvious data set, like wine Finally Documents.

#### 3.1 Blobs

#### 3.2 Moons

#### 3.3 Optimisers

#### 3.4 Batches

#### 3.5 Wine

#### 3.6 Rank Wiki Articles

### 4 Conclusion

---

## 5 Further Research

- Apply this to documents to get a sorted list.
- The "Quicksort" algorithm likely needs a random pivot to be efficient [9]

## 6 Text and References

Fractals are complex shapes that often occur from natural processes, in this report we hope to investigate the emergence of patterns and complex structures from natural phenomena. We begin with an investigation into fractals and the concept of dimension and then discuss links between fractal patterns and natural processes.

This is a Reference [10] and another [8] and yet another [1].

## 7 Fractals

Images are shown in figure 1.

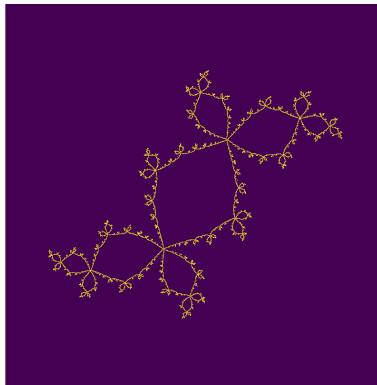


Figure 1: This is a test image showing the outline of a Julia set

## References

- [1] Christopher Burges. *From RankNet to LambdaRank to LambdaMART: An Overview* (MSR-TR-2010-82. Jan. 1, 2010. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2016/02/MSR-TR-2010-82.pdf> (cit. on p. 3).
- [2] Christopher Burges. *RankNet: A Ranking Retrospective*. Microsoft Research. July 7, 2015. URL: <https://www.microsoft.com/en-us/research/blog/ranknet-a-ranking-retrospective/> (visited on 02/15/2021) (cit. on p. 2).
- [3] Norbert Fuhr. “Optimum Polynomial Retrieval Functions Based on the Probability Ranking Principle”. In: *ACM Transactions on Information Systems* 7.3 (July 1, 1989), pp. 183–204. ISSN: 1046-8188. DOI: [10.1145/65943.65944](https://doi.org/10.1145/65943.65944). URL: <https://doi.org/10.1145/65943.65944> (visited on 02/15/2021) (cit. on p. 2).
- [4] Norbert Fuhr. “Probabilistic Models in Information Retrieval”. In: *The Computer Journal* 35.3 (June 1, 1992), pp. 243–255. ISSN: 0010-4620. DOI: [10.1093/comjnl/35.3.243](https://doi.org/10.1093/comjnl/35.3.243). URL: <https://doi.org/10.1093/comjnl/35.3.243> (visited on 02/15/2021) (cit. on p. 2).
- [5] Fredric C. Gey. “Inferring Probability of Relevance Using the Method of Logistic Regression”. In: *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR ’94. Berlin, Heidelberg: Springer-Verlag, Aug. 1, 1994, pp. 222–231. ISBN: 978-0-387-19889-7 (cit. on p. 2).
- [6] Tie-Yan Liu. “Learning to Rank for Information Retrieval”. In: *Foundations and Trends® in Information Retrieval* 3.3 (June 26, 2009), pp. 225–331. ISSN: 1554-0669, 1554-0677. DOI: [10.1561/15000000016](https://doi.org/10.1561/15000000016). URL: <https://www.nowpublishers.com/article/Details/INR-016> (visited on 02/15/2021) (cit. on p. 2).
- [7] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. New York: Cambridge University Press, 2008. 482 pp. ISBN: 978-0-521-86571-5 (cit. on p. 2).
- [8] Olympia Nicodemi, Melissa A. Sutherland, and Gary W. Towsley. *An Introduction to Abstract Algebra with Notes to the Future Teacher*. Upper Saddle River, NJ: Pearson Prentice Hall, 2007. 436 pp. ISBN: 978-0-13-101963-8 (cit. on p. 3).
- [9] Tim Roughgarden, director. *Quicksort Overview*. Jan. 28, 2017. URL: [https://www.youtube.com/watch?v=ETolcpLN7kk&list=PLEAYkSg4uSQ37A6\\_NrUnTHEKp6EkAxTma&index=25](https://www.youtube.com/watch?v=ETolcpLN7kk&list=PLEAYkSg4uSQ37A6_NrUnTHEKp6EkAxTma&index=25) (visited on 02/15/2021) (cit. on p. 3).
- [10] Enmei Tu et al. *A Graph-Based Semi-Supervised k Nearest-Neighbor Method for Nonlinear Manifold Distributed Data Classification*. June 3, 2016. arXiv: [1606.00985](https://arxiv.org/abs/1606.00985) [cs, stat]. URL: <http://arxiv.org/abs/1606.00985> (visited on 12/02/2020) (cit. on p. 3).
- [11] S. K.M. Wong and Y. Y. Yao. “Linear Structure in Information Retrieval”. In: *Proceedings of the 11th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR ’88. New York, NY, USA: Association for Computing Machinery, May 1, 1988, pp. 219–232. ISBN: 978-2-7061-0309-4. DOI: [10.1145/62437.62452](https://doi.org/10.1145/62437.62452). URL: <https://doi.org/10.1145/62437.62452> (visited on 02/14/2021) (cit. on p. 2).