PROJECT REPORT

Grouping Website News on Similarity Link
Using Page-rank Algorithm

Brillian Mahayana Putra
11.02.0053
2014/2015

FACULTY OF COMPUTER SCIENCE
SOEGIAPRANATA CATHOLIC UNIVERSITY
Jl. Pawiyatan Luhur IV/1, Bendan Duwur, SEMARANG 50234
Telp. 024-8441555 (hunting) Web: http://www.unika.ac.id
http://ikomunika.web.id/
APPROVAL AND RATIFICATION PAGE

PROJECT REPORT

Grouping Website News on Similarity Using Page-rank Algorithm

by

11.02.0053 – Brillant Mahayana Putra

This project report has been approved and ratified by the Dean of Faculty

of Computer Science and Supervisor on 15 July 2015

With approval,

Examiners,

Shinta Estra Wahyuningrum, S.Si., M.Cs
NPP: 058.1.2007.272

Hironimus Leong, S.Kom., M.Kom
NPP: 058.1.2007.273

Examiners,

Suyanto Edward Antonius, Jr., M.Sc

Rosita Herawati, ST., MIT
NPP: 058.1.2004.263

Supervisor,

Dean of Faculty of Computer Science,

Hironimus Leong, S.Kom., M.Kom
NPP: 058.1.2007.273
STATEMENT OF ORIGINALITY

I, the undersigned:

Name : Brillian Mahayana Putra
ID : 11.02.0053

Certify that this project was made by myself and not copy or plagiarize from other people, except that in writing expressed to the other article. If it is proven that this project was plagiarizes or copy the other, I am ready to accept a sanction.

Semarang, 15 July 2015
Brillian Mahayana Putra
11.02.0053
FOREWORD

First at all, give thanks to God for helping me and give me chance to finish this project as a graduation requirement. Lots of error were found during making the project. This project will not be completed without the support of everyone.

In making this project I would like to thank the Lord Jesus who always blessed me, my parents were always support me, Juliana Margaretha who always encourage and support me, and Mr. Hironimus Leong, S. Kom, M Kom who always guide and give me an ideas.
PREFACE

This project is Grouping Website News on Similarity Link Using Pagerank Algorithm. This program used to find incoming links in every news article. In making the project report is divided into several chapters.

In chapter 1, describes the background, scope, and purpose of the making of the program. So the program that created it has certain limitations. And the program that created will hopefully be a useful

Literature study also added on chapter 2 in this report. To make this project required knowledge is based on theories that already exists. With the support of existing theories. It is hoped the project will be made in accordance with the application of existing theory.

In project, planning is also very necessary such as in chapter 3. Project always have steps. Expected with the planning. The project will have a clear stage in the manufacturing process. Estimate time of completion of the program are also desperately needed. So there is an estimate of when the project is completed.

In chapter 4, explained about the steps project. In this chapter will describe in detail about the steps program. With a detail explanation hopefully make it easier for the readers to understand the content of this project.

Testing in project is needed to ensure that the program is already running properly. This stage is very important, because it can be known whether there is still an error in the program.

At the end of the project there will be a conclusion that can be drawn from this project. Although there is the conclusion of the program does not cover possible need for further development of this project.
ABSTRACT

Abstract — Grouping Website News on Similarity link using Page-rank Algorithm is classifying data based on similarity. With this grouping, user can easily know which website that contains a lot of news from a website news.

Page-rank algorithm is an algorithm that are rarely used in data mining. Page-rank will find where the links come from. First, Page-rank algorithm will search incoming link each news article then the incoming link will stored in array. After that incoming link will be group by similarity. So, links that are already grouped will produce a recommendation.

Final results of this grouping process is a recommendation for the reader. The reader can know any links that has a high degree of similarity links which linking in on the each article link.

Keywords : Data Mining, Two-Dimensional Array, Algorithm, Page-rank.
# Table of Contents

FOREWORD ......................................................................................................................... iv  
PREFACE ............................................................................................................................ v 
ABSTRACT............................................................................................................................ vi  
Chapter I : Introduction .............................................................................................................. 1  
1.1. Background .................................................................................................................. 1  
1.2. Scope .......................................................................................................................... 1  
1.3. Objective ..................................................................................................................... 1  
Chapter II : Literature Study .................................................................................................. 2  
2.1. Data Mining ............................................................................................................... 2  
2.2. Page-rank Algorithm .................................................................................................. 2  
2.3. Example ..................................................................................................................... 3  
2.4. Data Structure: Two Dimensional Array ..................................................................... 6  
Chapter III : Planning ......................................................................................................... 8  
3.1. Research Metodology .................................................................................................. 8  
3.2. Project Management ................................................................................................... 9  
Chapter IV : Analysis and Design ....................................................................................... 10  
4.1. Analysis ..................................................................................................................... 10  
4.1.2. Flow Chart ......................................................................................................... 10  
Chapter V : Implementation and Testing ........................................................................... 12  
5.1. Implementation .......................................................................................................... 12  
5.2. Testing ...................................................................................................................... 22  
Chapter VI : Conclusion ..................................................................................................... 28  
6.1. Conclusion .................................................................................................................. 28  
6.2. Further Research ....................................................................................................... 28  
References .......................................................................................................................... 29  
APPENDIX ....................................................................................................................... 30
**Table of Figure**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Example Two Dimensional Array</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Flow chart</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Choose the source</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>Start &amp; End Date</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>Input the Source</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>Input Date</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>Result Links Articles</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>Get Linking In All Articles</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>Sorting by Total</td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>Grouping Done</td>
<td>26</td>
</tr>
<tr>
<td>11</td>
<td>Get Recommendation</td>
<td>26</td>
</tr>
</tbody>
</table>
Table of Tables

Table 1. Example of Grouping ................................................................. 5
Table 2. Project Management ................................................................. 9