Project Proposal for I524

ABHISHEK GUPTA[1,*] **AND AVADHOOT AGASTI**[1,**]

[1] *School of Informatics and Computing, Bloomington, IN 47408, U.S.A.*
[*] *Corresponding authors: abhigupt@iu.edu*
[**] *Corresponding authors: aagasti@iu.edu*

project-1: Data mining for a wiki url , March 5, 2017

Keywords: Cloud, I524

https://github.com/cloudmesh/classes/blob/master/docs/source/format/report/report.pdf

1. PROBLEM

Given a wiki URL of a person, find out his details like School, Spouse, Coaches, language, alma-meter etc Typically, the wiki page has all this information available but in the free form text. We need to converting it into structured data format so that it can help us analyze the people, from the networks etc We can create a network by navigating the people mentioned in the wiki page.

2. SOLUTION

Use tensor flow [1] to create word vectors. Train it using manual tagging and then use the model for analytics and prediction.

Technology Name	Purpose
tensor flow [2]	work vector and model
spark [3]	data analysis
ansible [4]	automated deployment
python [1]	development

3. DEPLOYMENT

Solution will be deployed using Ansible [4] playbook.

REFERENCES

[1] "Tensor Python, python tensor flow development library," Web Page, accessed: 2017-02-26. [Online]. Available: https://www.tensorflow.org/api_docs/python/

[2] "Tensor Flow, an open-source software library for machine intelligence," Web Page, accessed: 2017-02-26. [Online]. Available: https://www.tensorflow.org/get_started/get_started

[3] "Tensor Spark, distributed tensorflow on spark," Web Page, accessed: 2017-02-26. [Online]. Available: https://github.com/adatao/tensorspark

[4] "Ansible, deploy apps. manage systems. crush complexity," Web Page, accessed: 2017-02-26. [Online]. Available: https://www.ansible.com/

6. Conclusion Using this wiki analysis we should be able to build a network based on wiki data.

Acknowledgement We acknowledge our professor Gregor von Laszewski and all associate instructors for helping us and guiding us throughout this project.