**Capstone – Project synopsis\_aimlChnOct19Grp4**

**Problem Statement:**

To build a Machine Learning model for Custom NER using deep neural network to deal with information extraction from the real-world data and to build a classifier based on polarity of the movie reviews. In order to get insightful information, we approach the machine and deep learning techniques for sentiment analysis and entity recognition over labelled and un processed data.

**Literature Survey:**

**Sample Data:** Movie Review Data

Movie Review Data extracted for the use of sentiment-analysis experiments which was orchestrated on 2002 by Cornell University, New York. Collections of movie-review documents labelled with respect to their overall

* Sentiment polarity datasets (positive or negative)
* Sentiment scale datasets
* Subjectivity datasets
* Data Source: https://www.cs.cornell.edu/people/pabo/movie-review-data/

**Tentative List of Algorithms:**

**References:**

[1] Bo Pang, Lillian Lee and Shivakumar Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques, Proceedings of EMNLP 2002”.

[2] Bo Pang and Lillian Lee, “[A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts](http://www.cs.cornell.edu/home/llee/papers/cutsent.home.html), Proceedings of ACL 2004”.

[3] Bo Pang and Lillian Lee, “[Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales](http://www.cs.cornell.edu/home/llee/papers/pang-lee-stars.home.html), Proceedings of ACL 2005”.

[4] Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts, “Learning Word Vectors for Sentiment Analysis”

[5] Jorge Carrillo de Albornoz, Laura Plaza, Pablo Gervás, “A Hybrid Approach to Emotional Sentence Polarity and Intensity Classification“

[6] Bo Pang and Lillian Lee, “A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts”

[7] Ainur Yessenalina, Yisong Yue & Claire Cardie, “Multi-level Structured Models for Document-level Sentiment Classification”

[8] Rudy Prabowo1 , Mike Thelwall, “Sentiment Analysis: A Combined Approach”

[9] Chenghua Lin, Yulan He, Richard Everson, “A Comparative Study of Bayesian Models for Unsupervised Sentiment Detection”

[10] Yulan He Chenghua Lin† Harith Alani, “Automatically Extracting Polarity-Bearing Topics for Cross-Domain Sentiment Classification”

[11] Richard Socher, Jeffrey Pennington∗, Eric H. Huang, Andrew Y. Ng, Christopher D. Manning,” Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions”

[12] V. Suresh, Ashok Veilumuthu, Avanthi Krishnamurthy, “A Non-syntactic Approach for Text Sentiment Classification with Stopwords”

[13] Amanda Hutton, Alexander Liu, Cheryl Martin, “Crowdsourcing Evaluations of Classifier Interpretability”

[14] Seungyeon Kim, Fuxin Li, Guy Lebanon, and Irfan Essa, “Beyond Sentiment: The Manifold of Human Emotions”

[15] Alexander Erdmann, David Joseph Wrisley, Benjamin Allen, Christopher Brown, Sophie Cohen-Boden´ es, ` ∗ Micha Elsner, Yukun Feng, Brian Joseph, Beatrice Joyeux-Prunel ´ ∗ and Marie-Catherine de Marneffe, “Practical, Efficient, and Customizable Active Learning for Named Entity Recognition in the Digital Humanities”

[16] Łukasz Augustyniak, Piotr Szyma ´nski, Tomasz Kajdanowicz and Włodzimierz Tuligłowicz, “Comprehensive Study on Lexicon-based Ensemble Classification Sentiment Analysis”

[17] Justin Martineau, and Tim Finin, “Delta TFIDF: An Improved Feature Space for Sentiment Analysis”

[18] Matthew E. Peters , Mark Neumann , Mohit Iyyer , Matt Gardner, Christopher Clark∗ , Kenton Lee∗ , Luke Zettlemoyer, “Deep contextualized word representations”

[19] Nora Hollenstein, Ce Zhang, “Entity Recognition at First Sight: Improving NER with Eye Movement Information”

[20] Georgios Paltoglou, Mike Thelwall, “A study of Information Retrieval weighting schemes for sentiment analysis”

[21] Tao Gui1 , Ruotian Ma , Qi Zhang , Lujun Zhao, Yu-Gang Jiang and Xuanjing Huang, “CNN-Based Chinese NER with Lexicon Rethinking”

[22] Xuezhe Ma and Eduard Hovy, “End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF”

[23] Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, Chris Dyer, Carnegie Mellon University, NLP Group, Pompeu Fabra University, “Neural Architectures for Named Entity Recognition”

[24] Hui Chen , Zijia Lin , Guiguang Ding , Jianguang Lou, Yusen Zhang, Borje Karlsson, “GRN: Gated Relation Network to Enhance Convolutional Neural Network for Named Entity Recognition∗”

[25] Zihan Wang∗, Jingbo Shang∗, Liyuan Liu∗, Lihao Lu, Jiacheng Liu, Jiawei Han, “CrossWeigh: Training Named Entity Tagger from Imperfect Annotations”

[26] Stephen Mayhew, Nitish Gupta, Dan Roth, “Robust Named Entity Recognition with Truecasing Pretraining”

[27] Hongyu Lin , Yaojie Lu, Jialong Tang, Xianpei Han, Le Sun, Zhicheng Wei, Nicholas Jing Yuan4, “A Rigorous Study on Named Entity Recognition: Can Fine-tuning Pretrained Model Lead to the Promised Land?”

[28] Pierre Lison and Aliaksandr Hubin, Jeremy Barnes and Samia Touileb, “Named Entity Recognition without Labelled Data: A Weak Supervision Approach”

[29] Prem Melville and Raymond J. Mooney and Ramadass Nagarajan, “Content-Boosted Collaborative Filtering for Improved Recommendations”

[30] Sangeeta Oswal, Ravikumar Soni, Omkar Narvekar, Abhijit Pradha, “Named Entity Recognition and Aspect based Sentiment Analysis”