of prediction. Based on this phenomenon, we chose LDA model, which is both well-performed and stable, to do feature selection and got the result as shown in Table 6.

Table 6. Feature Selection based on LDA model

Run	Composition	AvgCosin	AvgAbsLoss
4	baseline	0.8812	0.1380
5	baseline+f7	0.8825	0.1343
6	baseline+f7+f20	0.8831	0.1339
7	baseline+f7+f14	0.8841	0.1335
8	baseline+f7+f14+f13	**0.8886**	**0.1286**

6. CONCLUSIONS

This paper details the approach DUT-CH group addressed for Temporalia task at the NTCIR-12. We participate in TID subtask (Chinese) aiming at predicting the distribution of four temporal intents. For TID Chinese subtask, the formal run results were produced by adopting all the designed features to linear SVC model, Logistic Regression model with $l1$ penalty and Random Forest model with class weight balanced. After the submission of the formal run, we did further experiments to compare different models and feature composition and finally got a better and more stable result by LDA model and selected time gap, word-based probability distribution vector, temporal trigger word, Google Trends' time gap, center word and its Part-of-speech as good features.

In the future work, we will exploit more implicit features by analyzing time-series data, Google Trends and the documents returned by search engines like Google Search and Baidu Search.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] R. Campos, G. Dias, AM. Jorge and A. Jatowt. Survey of temporal information retrieval and related applications. *ACM Computing Surveys (CSUR)*. 47(2):15, 2015.

[2] K. Nattiya, R. Blanco, and K. Nørvåg. Temporal Information Retrieval. *Foundations and Trends in Information Retrieval*, 9(2): 91-208, 2015.

[3] H. Joho, A. Jatowt, and B. Roi. A Survey of Temporal Web Search Experience. In *Proceedings of the 22nd International Conference on World Wide Web (Companion)*. pages 1101–1108, 2013

[4] R. Jones and F. Diaz. Temporal profiles of queries. *ACM Transactions on Information Systems (TOIS)*, 25(3): 14, 2007.

[5] Vlachos, M., Meek, C., Vagena, Z., Gunopulos, D.: Identifying similarities, periodicities and bursts for online search queries. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pages 131-142, 2004

[6] A. Kulkarni, J. Teevan, K. M. Svore and S. T. Dumais. Understanding temporal query dynamics. In *Proceedings of*

the fourth ACM international conference on Web search and data mining*, pages 167-176, 2011

[7] P. Ren, Z. Chen, X. Song, B. Li, H. Yang and J. Ma. Understanding temporal intent of user query based on time-based query classification. In: *Natural Language Processing and Chinese Computing*, pages 334-345, 2013

[8] N. Parikh and N. Sundaresan. Scalable and near real-time burst detection from ecommerce queries. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 972-980, 2008

[9] N. Kanhabua and K. Nørvåg. Exploiting time-based synonyms in searching document archives. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 79-88, 2010

[10] H. Joho, A. Jatowt, R. Blanco, H. Yu, and S. Yamamoto. Overview of NTCIR-12 temporal information access (temporalia-2) task. In *Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies*, 2016.

[11] A. Spink, D. Wolfram, MBJ. Jansen, and T. Saracevic, Searching the web: The public and their queries. *Journal of the American society for information science and technology*, 52(3): 226-234, 2001.

[12] H. Joho, A. Jatowt, R. Blanco, H. Naka and S. Yamamoto. Overview of NTCIR-11 Temporal Information Access (Temporalia) Task. In *Proceedings of the 11th NTCIR conference*, pages 429-437, 2014.

[13] C. D. Manning, M. Surdeanu, J. Bauer, J. R. Finkel, S. Bethard, and D. McClosky. The Stanford CoreNLP Natural Language Processing Toolkit. In *ACL (System Demonstrations)*, pages 55-60, 2014.

[14] H. Li, J. Strötgen, J. Zell and M. Gertz. Chinese Temporal Tagging with HeidelTime. In *EACL*, pages 133-137, 2014.

[15] H. Yu, X. Kang, and F. Ren. TUTA1 at the NTCIR-11 Temporalia Task. In *Proceedings of the 11th NTCIR Conference*, pages 461-467, 2014.

[16] T. Mikolov, K. Chen, G. Corrado and J. Dean. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.

[17] W. McKinney, J. Perktold and S. Seabold. Time series analysis in Python with statsmodels. *Jarrodmillman. Com*, pages 96-102, 2011