



School of Computing, Engineering, and the Built Environment Edinburgh Napier University

PHD STUDENT PROJECT

Funding and application details

Funding status: Fully funded project (worldwide)

Application instructions:

Detailed instructions are available at <https://blogs.napier.ac.uk/scebe-research/available-phd-student-projects/>

Prospective candidates are encouraged to contact the Director of Studies (see details below) to discuss the project and their suitability for it.

Project details

Supervisory Team:

- DIRECTOR OF STUDY: Peter J Barclay (Email: P.Barclay@napier.ac.uk)
- 2ND SUPERVISOR: TBC

Subject Group: Computer science

Research Areas: Artificial Intelligence, Data Science, Machine Learning

Project Title: Characterising Automatically Generated Text

Project description:

Research is needed to understand the differences between human and machine generated text. Prior research has focused on identification of identification of deceptive text, such as phishing attempts or bot-generated tweets. There is, however, less work on identifying other forms of generated text such as automatic translations, or text that has been recorded by 'essay assistant' software. This project would focus on characterising the differences between human generated and machine generated text, for example by comparing originally authored material with automatically rewritten versions.

References:

- [1] Afroz, S., Brennan, M., & Greenstadt, R. (2012). Detecting hoaxes, frauds, and deception in writing style online. 2012 IEEE Symposium on Security and Privacy, 461–475.
- [2] Boudin, F., Mougard, H., & Cram, D. (2016). How document pre-processing affects keyphrase extraction performance. ArXiv Preprint ArXiv:1610.07809.
- [3] Dou, Y., Forbes, M., Koncel-Kedziorski, R., Smith, N., & Choi, Y. (2022). Is GPT-3 Text Indistinguishable from Human Text? Scarecrow: A Framework for Scrutinizing Machine Text. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 7250–7274. <https://doi.org/10.18653/v1/2022.acl-long.501>
- [4] Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45(1), 1–23.
- [5] Ippolito, D., Duckworth, D., Callison-Burch, C., & Eck, D. (2019). Automatic detection of generated text is easiest when humans are fooled. ArXiv Preprint ArXiv:1911.00650.
- [6] Jawahar, G., Abdul-Mageed, M., & Lakshmanan, L. V. (2020). Automatic detection of machine generated text: A critical survey. ArXiv Preprint ArXiv:2011.01314.
- [7] Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5), 665–675.
- [8] Varshney, L. R., Keskar, N. S., & Socher, R. (2020). Limits of detecting text generated by large-scale language models. 2020 Information Theory and Applications Workshop (ITA), 1–5.

Candidate characteristics

Education:

A first-class honours degree, or a distinction at master level, or equivalent achievements in Computer Science, Data Science, Linguistics

Subject knowledge:

- Experience of fundamental machine learning models
- Competent in Python programming
- Knowledge of linguistics or natural language processing

Essential attributes:

- Good written and oral communication skills
- Strong motivation, with evidence of independent research skills relevant to the project
- Good time management

Desirable attributes: