國立高雄大學資訊工程學系碩士論文計畫書

**資訊工程學系 V.1**

|  |  |  |
| --- | --- | --- |
| 一、碩士論文計畫書規定：   1. 碩士生應於擬畢業學期開學一週內，提出碩士論文計畫書。 2. 指導教授變更時，須重新提出碩士論文計畫書。 | | |
| 二、碩士生資料 | 姓名：周函霖 | 學號：m1085504 |
| 三、研究題目：應用深度學習方法於程式碼轉換的快速驗證 | | |
| 研究動機與目的：  本研究利用自然語言生成模型 GPT-2[1]、MASS [2]和 BART [3]當作程式碼轉換模型來進行程式碼轉換作業。為了加快程式碼轉換後的驗證，本研究提出variational simhash演算法，比較範例程式與初階程式的相似度。除了提高程式碼相似度的準確性，同時也減少了合格程式的數量和編譯時間。還提出了piecewise longest common subsequence演算法，檢測執行結果的一致性。不僅減少了字串比較的次數，而且利用較少的記憶體空間暫時儲存檢測結果，使其有更多的記憶體進行字串比較運算，如此可以加速執行結果的一致性檢測。最後，使用local interpretable model-agnostic explanations[4]來解釋模型進行推論的決定。  時程安排：  上學期與指導教授討論題目、研究方法與框架，下學期開始建立實驗環境並進行實驗、投稿論文與碩士論文。  參考資料：   1. J. J. Bird, M. Pritchard, A, Fratini, A. Ekárt, D. R. Faria,” Synthetic Biological Signals Machine-Generated by GPT-2 Improve the Classification of EEG and EMG Through Data Augmentation,” IEEE Robotics and Automation Letters, Vol. 6, Issue: 2, pp. 3498 – 3504,April 2021. 2. K. Song, X. Tan, T. Qin, J. Lu and T.Y. Liu, “MASS: Masked Sequence to Sequence Pre-training for Language Generation,” International Conference on Machine Learning( ICML), Long Beach, CA , 2019. 3. M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, “Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension,” Annual Meeting of the Association for Computational Linguistics (ACL), Online, pp. 7871-7880, 2020 4. M. T. Ribeiro, S. Singh, C. Guestrin, “"Why Should I Trust You?": Explaining the Predictions of Any Classifier,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 1135–1144, 2016. | | |
| 五、碩士生簽名： 日期： | | |
| 六、指導教授簽名： 日期： | | |
| 系主任簽名：　　　　　　　　　　　　　　　　　　日期： | | |