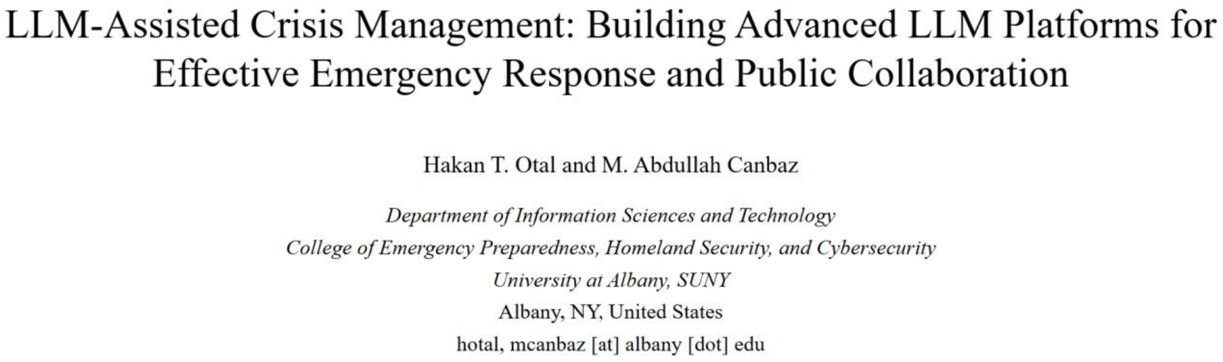
Project1 – LLM Assisted Crisis Management

Dataset link - https://crisisnlp.qcri.org/crisismmd

Alam, F., Ofli, F., & Imran, M. (2018, June). Crisismmd: Multimodal twitter datasets from natural disasters. In *Proceedings of the international AAAI conference on web and social media* (Vol. 12, No. 1).

Preference data – California wildfire

Related Paper –

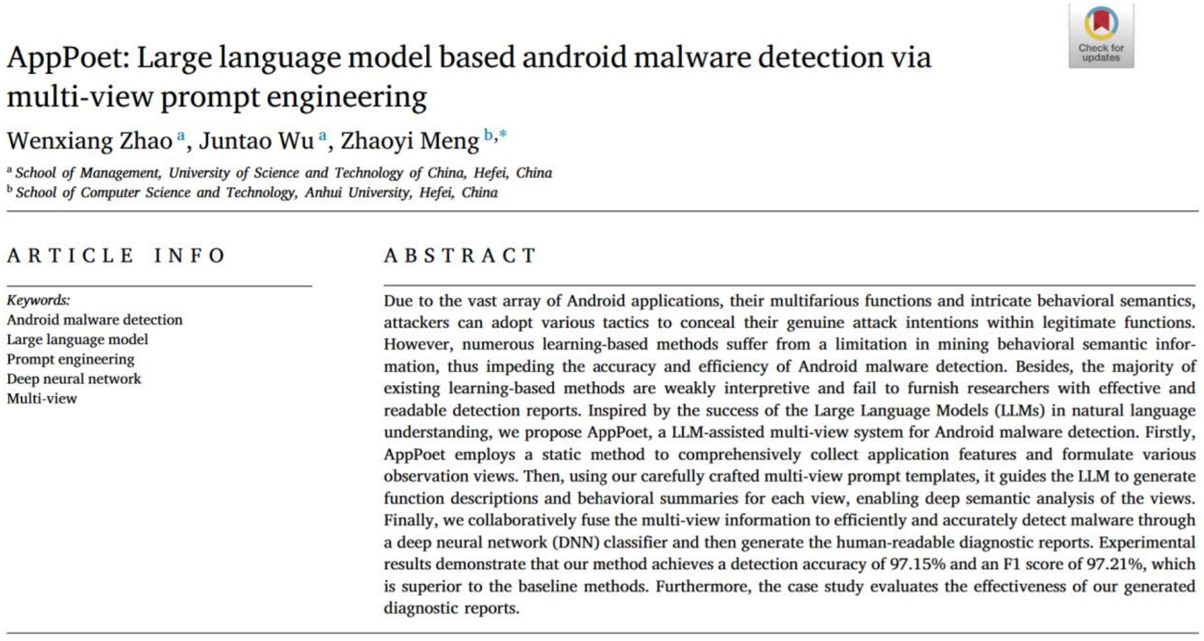


LLM-Assisted Crisis Management: Building Advanced LLM Platforms for Effective Emergency Response and Public Collaboration

Hakan T. Otal and M. Abdullah Canbaz

Department of Information Sciences and Technology
College of Emergency Preparedness, Homeland Security, and Cybersecurity
University at Albany, SUNY
Albany, NY, United States
hotal, mcanbaz [at] albany [dot] edu

Project2 - LLM based Ransomware/Malware/Spyware Detection



AppPoet: Large language model based android malware detection via multi-view prompt engineering

Wenxiang Zhao [a], Juntao Wu [a], Zhaoyi Meng [b,*]

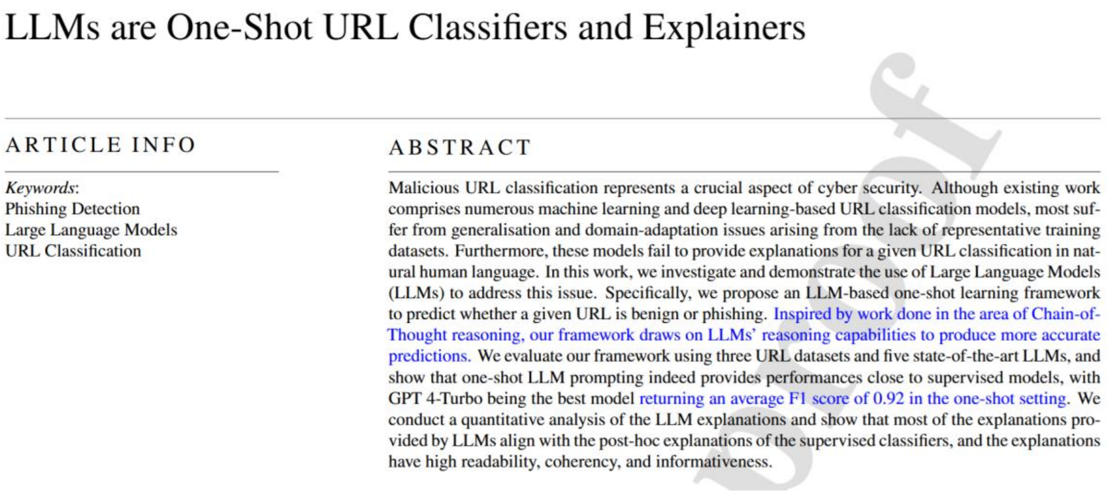
[a] *School of Management, University of Science and Technology of China, Hefei, China*
[b] *School of Computer Science and Technology, Anhui University, Hefei, China*

ARTICLE INFO

Keywords:
Android malware detection
Large language model
Prompt engineering
Deep neural network
Multi-view

ABSTRACT

Due to the vast array of Android applications, their multifarious functions and intricate behavioral semantics, attackers can adopt various tactics to conceal their genuine attack intentions within legitimate functions. However, numerous learning-based methods suffer from a limitation in mining behavioral semantic information, thus impeding the accuracy and efficiency of Android malware detection. Besides, the majority of existing learning-based methods are weakly interpretive and fail to furnish researchers with effective and readable detection reports. Inspired by the success of the Large Language Models (LLMs) in natural language understanding, we propose AppPoet, a LLM-assisted multi-view system for Android malware detection. Firstly, AppPoet employs a static method to comprehensively collect application features and formulate various observation views. Then, using our carefully crafted multi-view prompt templates, it guides the LLM to generate function descriptions and behavioral summaries for each view, enabling deep semantic analysis of the views. Finally, we collaboratively fuse the multi-view information to efficiently and accurately detect malware through a deep neural network (DNN) classifier and then generate the human-readable diagnostic reports. Experimental results demonstrate that our method achieves a detection accuracy of 97.15% and an F1 score of 97.21%, which is superior to the baseline methods. Furthermore, the case study evaluates the effectiveness of our generated diagnostic reports.



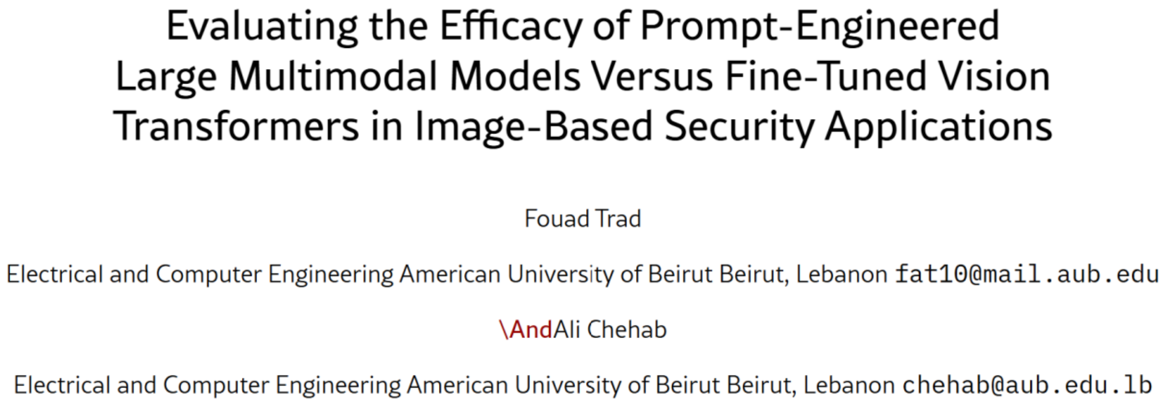
LLMs are One-Shot URL Classifiers and Explainers

ARTICLE INFO

Keywords:
Phishing Detection
Large Language Models
URL Classification

ABSTRACT

Malicious URL classification represents a crucial aspect of cyber security. Although existing work comprises numerous machine learning and deep learning-based URL classification models, most suffer from generalisation and domain-adaptation issues arising from the lack of representative training datasets. Furthermore, these models fail to provide explanations for a given URL classification in natural human language. In this work, we investigate and demonstrate the use of Large Language Models (LLMs) to address this issue. Specifically, we propose an LLM-based one-shot learning framework to predict whether a given URL is benign or phishing. Inspired by work done in the area of Chain-of-Thought reasoning, our framework draws on LLMs' reasoning capabilities to produce more accurate predictions. We evaluate our framework using three URL datasets and five state-of-the-art LLMs, and show that one-shot LLM prompting indeed provides performances close to supervised models, with GPT 4-Turbo being the best model returning an average F1 score of 0.92 in the one-shot setting. We conduct a quantitative analysis of the LLM explanations and show that most of the explanations provided by LLMs align with the post-hoc explanations of the supervised classifiers, and the explanations have high readability, coherency, and informativeness.



Evaluating the Efficacy of Prompt-Engineered Large Multimodal Models Versus Fine-Tuned Vision Transformers in Image-Based Security Applications

Fouad Trad

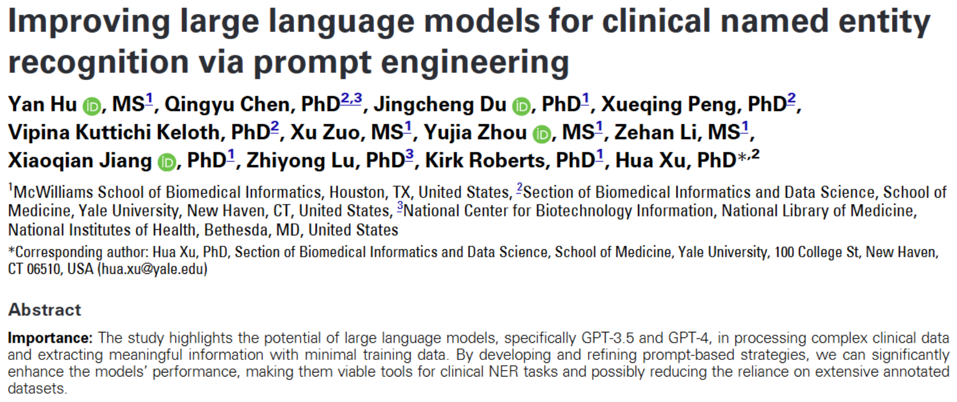
Electrical and Computer Engineering American University of Beirut Beirut, Lebanon fat10@mail.aub.edu

\AndAli Chehab

Electrical and Computer Engineering American University of Beirut Beirut, Lebanon chehab@aub.edu.lb

Project3 – LLMs to extract Relevant Information (NER) from Clinical Notes

Dataset links - https://huggingface.co/datasets/GBaker/MedQA-USMLE-4-options?row=0



Improving large language models for clinical named entity recognition via prompt engineering

Yan Hu, MS[1], Qingyu Chen, PhD[2,3], Jingcheng Du, PhD[1], Xueqing Peng, PhD[2], Vipina Kuttichi Keloth, PhD[2], Xu Zuo, MS[1], Yujia Zhou, MS[1], Zehan Li, MS[1], Xiaoqian Jiang, PhD[1], Zhiyong Lu, PhD[3], Kirk Roberts, PhD[1], Hua Xu, PhD*[,2]

[1]McWilliams School of Biomedical Informatics, Houston, TX, United States, [2]Section of Biomedical Informatics and Data Science, School of Medicine, Yale University, New Haven, CT, United States, [3]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, United States

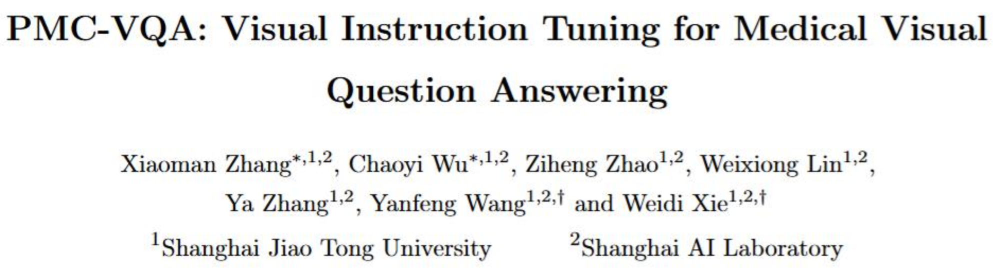
*Corresponding author: Hua Xu, PhD, Section of Biomedical Informatics and Data Science, School of Medicine, Yale University, 100 College St, New Haven, CT 06510, USA (hua.xu@yale.edu)

Abstract

Importance: The study highlights the potential of large language models, specifically GPT-3.5 and GPT-4, in processing complex clinical data and extracting meaningful information with minimal training data. By developing and refining prompt-based strategies, we can significantly enhance the models' performance, making them viable tools for clinical NER tasks and possibly reducing the reliance on extensive annotated datasets.

Project4 – Multimodal LLMs Medical Visual Question Answering

Dataset link - https://huggingface.co/datasets/xmcmic/PMC-VQA

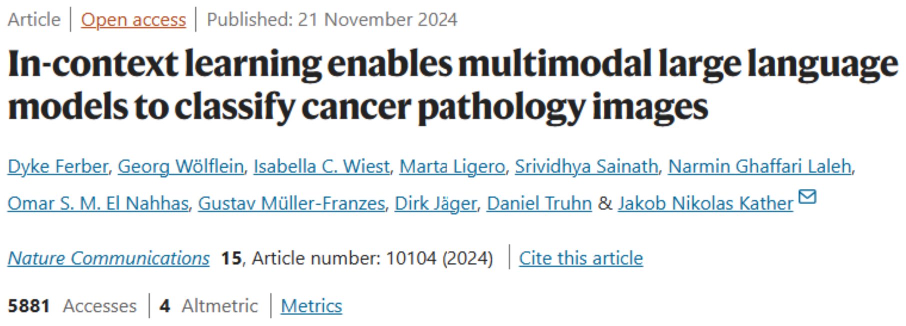


PMC-VQA: Visual Instruction Tuning for Medical Visual Question Answering

Xiaoman Zhang[*,1,2], Chaoyi Wu[*,1,2], Ziheng Zhao[1,2], Weixiong Lin[1,2], Ya Zhang[1,2], Yanfeng Wang[1,2,†] and Weidi Xie[1,2,†]

[1]Shanghai Jiao Tong University [2]Shanghai AI Laboratory

Project5 – Multimodal LLMs to Classify Cancer Pathology Images



Article | Open access | Published: 21 November 2024

In-context learning enables multimodal large language models to classify cancer pathology images

Dyke Ferber, Georg Wölflein, Isabella C. Wiest, Marta Ligero, Srividhya Sainath, Narmin Ghaffari Laleh, Omar S. M. El Nahhas, Gustav Müller-Franzes, Dirk Jäger, Daniel Truhn & Jakob Nikolas Kather ✉

Nature Communications **15**, Article number: 10104 (2024) | Cite this article

5881 Accesses | **4** Altmetric | Metrics

Dataset links –

https://huggingface.co/datasets/1aurent/PatchCamelyon

https://huggingface.co/datasets/mamunrobi35/mhist_binary

Project6 – LLMs to Extract/Filter Information from Structured Tabular Dataset

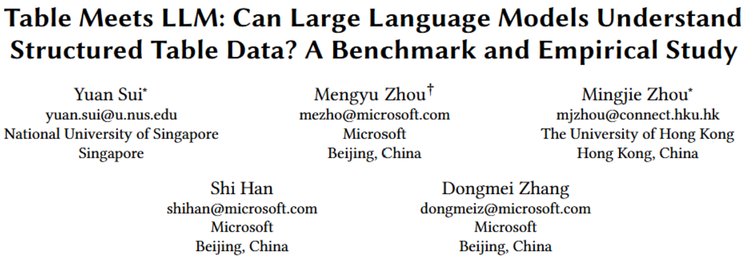


Table Meets LLM: Can Large Language Models Understand Structured Table Data? A Benchmark and Empirical Study

Yuan Sui[*]
yuan.sui@u.nus.edu
National University of Singapore
Singapore

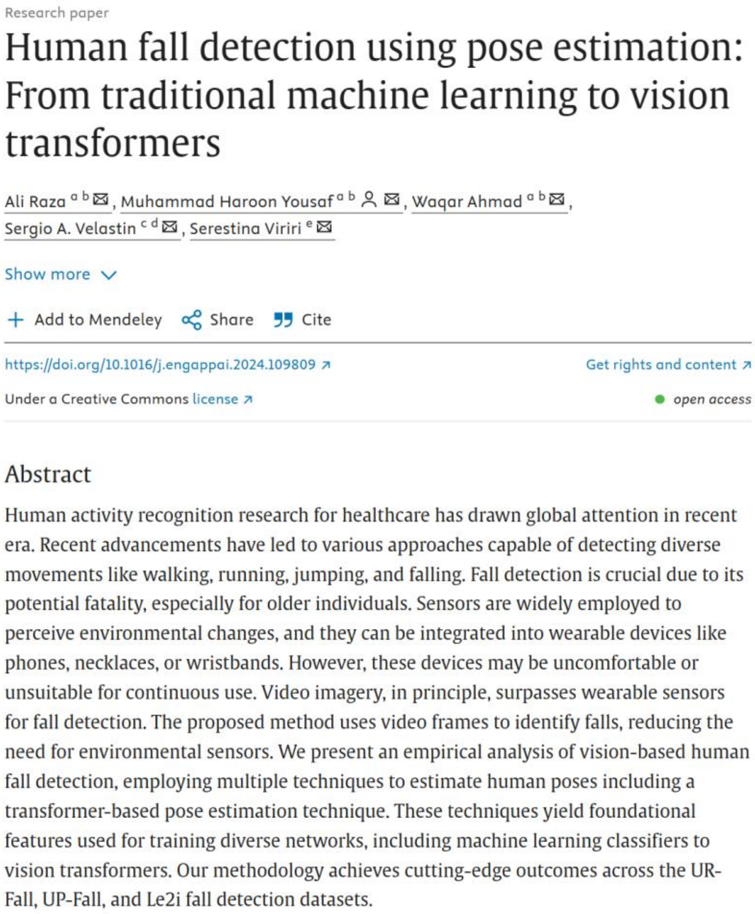
Mengyu Zhou[†]
mezho@microsoft.com
Microsoft
Beijing, China

Mingjie Zhou[*]
mjzhou@connect.hku.hk
The University of Hong Kong
Hong Kong, China

Shi Han
shihan@microsoft.com
Microsoft
Beijing, China

Dongmei Zhang
dongmeiz@microsoft.com
Microsoft
Beijing, China

Project7 – Multimodal (Vision) LLMs for Fall Detection



Research paper

Human fall detection using pose estimation: From traditional machine learning to vision transformers

Ali Raza [a b], Muhammad Haroon Yousaf [a b], Waqar Ahmad [a b], Sergio A. Velastin [c d], Serestina Viriri [e]

https://doi.org/10.1016/j.engappai.2024.109809

● open access

Abstract

Human activity recognition research for healthcare has drawn global attention in recent era. Recent advancements have led to various approaches capable of detecting diverse movements like walking, running, jumping, and falling. Fall detection is crucial due to its potential fatality, especially for older individuals. Sensors are widely employed to perceive environmental changes, and they can be integrated into wearable devices like phones, necklaces, or wristbands. However, these devices may be uncomfortable or unsuitable for continuous use. Video imagery, in principle, surpasses wearable sensors for fall detection. The proposed method uses video frames to identify falls, reducing the need for environmental sensors. We present an empirical analysis of vision-based human fall detection, employing multiple techniques to estimate human poses including a transformer-based pose estimation technique. These techniques yield foundational features used for training diverse networks, including machine learning classifiers to vision transformers. Our methodology achieves cutting-edge outcomes across the UR-Fall, UP-Fall, and Le2i fall detection datasets.