

Tsukuba x NTPU

Bilateral Academic Exchange Program





Generative AI for ESG Data Analytics and Sustainability Innovation



Min-Yuh Day, Ph.D.

Professor, <u>Institute of information Management</u>
Director, Fintech and Green Finance Center (FGFC)

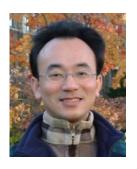
Division Director, Sustainable Development, Sustainability Office

National Taipei University

https://web.ntpu.edu.tw/~myday







Professor Min-Yuh Day



2020 Cohort



Accredited Educator





Professor, Information Management, NTPU

Director, Intelligent Financial Innovation Technology, IFIT Lab, IM, NTPU

Director, Fintech and Green Finance Center (FGFC), NTPU

Division Director, Sustainable Development, Sustainability Office, NTPU

Visiting Scholar, IIS, Academia Sinica

Ph.D., Information Management, NTU

Publications Co-Chairs, International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013-)

Program Co-Chair, IEEE International Workshop on Empirical Methods for Recognizing Inference in TExt (IEEE EM-RITE 2012-)

Publications Chair, The IEEE International Conference on Information Reuse and Integration for Data Science (IEEE IRI 2007-)









Outline



- Overview of ESG
 - Environmental, Social, Governance
- Challenges in ESG Data Analytics
 - Fragmented, Unstructured, Diverse
- Generative AI as a Solution
 - Synthesis, Reporting, and Prediction
- Sustainability Innovation

Evolution of Sustainable Finance Research





Sustainable Development Goals (SDGs)







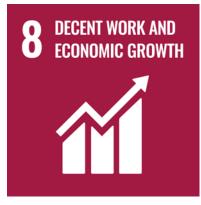
































Sustainable Development Goals (SDGs) and 5P



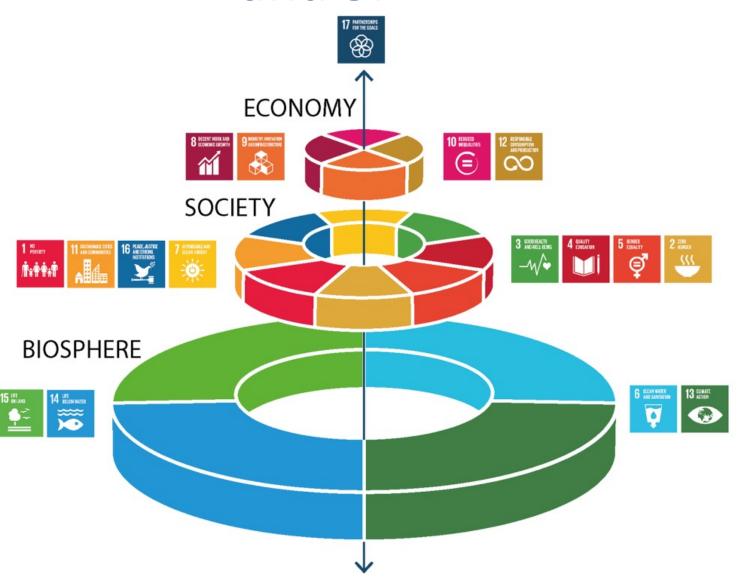
Partnership

Peace

Prosperity

People

Planet



ESG to 17 SDGs



ENVIRONMENT



14 LIFE BELOW WATER





13 CLIMATE ACTION





SOCIAL













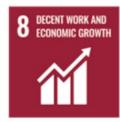






GOVERNANCE











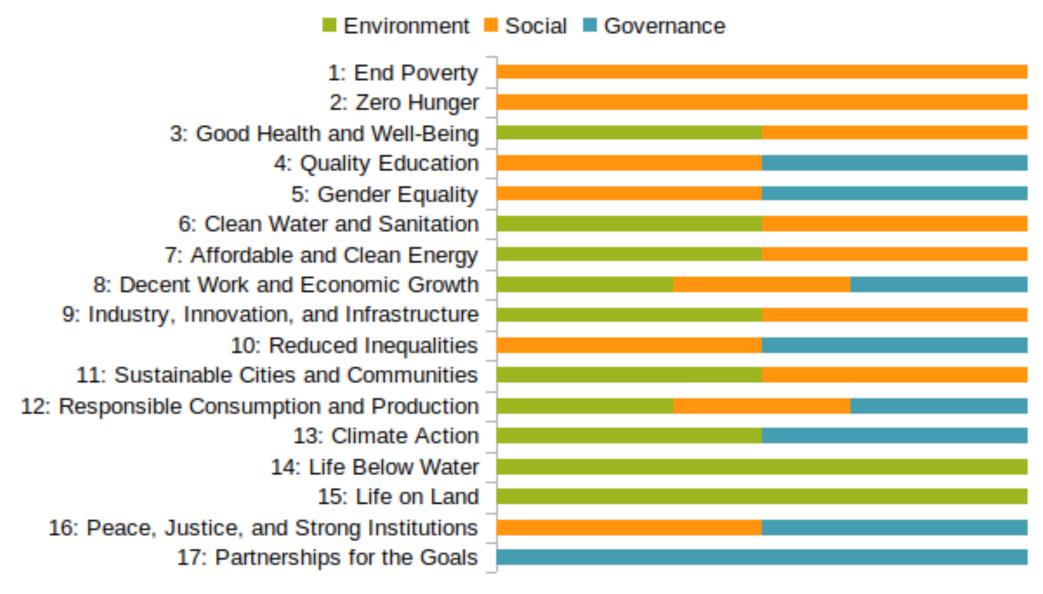






ESG to 17 SDGs





Net-Zero Transformation



Ambition

 Aligned to achieving global net zero by no later than 2050 & to limit warming to 1.5° C

Governance

Accountability driven from the top

Strategy

Embedded and aligned net zero into company strategy

Enterprise

 Key operating model changes in support of transformation

Supply chains

Transformed net zero supply chains

Innovation

 Developed innovation and technologies to deliver net zero

Finance

 Financing the net zero transformation

Transparency

Communicating action

Engagement

Enhancing the pace and scale of net zero action

ESG Challenges and Opportunities



Challenges

- Fragmented and unstructured ESG data.
- Lack of standardization and transparency.
- Timeliness of data availability.

Opportunities

- Rising demand for actionable ESG insights.
- Innovation in sustainable solutions and policies.
- Generative AI as a tool for transformation.



Generative Al Powering Digital Sustainability Transformation



Generative Al-Driven ESG Report Generation Technology

Industrial Technology Research Institute (ITRI), Fintech and Green Finance Center (FGFC, NTPU), NTPU-113A513E01, 2024/03/01~2024/12/31

Sustainability and ESG Data Analytics





Generative AI for ESG Data Analytics



- Data Integration and Enrichment:
 - Synthesizing structured and unstructured ESG data.
- Automated Reporting and Insight Generation:
 - Tailored ESG reports and insights for stakeholders.
- Scenario Modeling and Forecasting:
 - Simulating potential risks and opportunities.
- Addressing Bias and Ensuring Accountability:
 - Transparent, fair, and ethical AI deployment.

Generative AI and LLMs for Sustainability and ESG Data Analytics





Sustainability Innovation with Generative Al



- Sustainable Product Design:
 - Eco-friendly designs minimizing waste and energy.
- Policy Formulation and Implementation:
 - Al-driven simulations for effective policies.
- Stakeholder Engagement and Awareness:
 - Communicating ESG strategies with compelling Al-driven visuals.

Generative AI for ESG Rating and Reporting Generation





Future Directions



- Integrating blockchain, IoT, and digital twins.
- Democratizing AI tools for all stakeholders.
- Promoting collaboration among experts and communities.

Conclusion



- Generative AI is transforming ESG analytics and sustainability innovation.
- Collaboration among researchers, policymakers, and innovators is key.
- Generative AI to build a sustainable future.



2022







IMNTPU at the NTCIR-16 FinNum-3 Task: Data Augmentation for Financial Numclaim Classification

¹ Information Management, National Taipei University, New Taipei City, Taiwan ² Zeals Co., Ltd. Tokyo, Japan



Yung-Wei Teng ¹



Pei-Tz Chiu¹



Ting-Yun Hsiao ¹



Mike Tian-Jian Jiang ² Min-Yuh Day ^{1,*}





2022







IMNTPU Dialogue System Evaluation at the NTCIR-16 DialEval-2 **Dialogue Quality and Nugget Detection**

¹ Information Management, National Taipei University, New Taipei City, Taiwan ² Zeals Co., Ltd. Tokyo, Japan



Ting-Yun Hsiao ¹



Yung-Wei Teng ¹



Pei-Tz Chiu ¹



Mike Tian-Jian Jiang ² Min-Yuh Day ^{1,*}







NTCIR-16

FinNum-3

IMNTPU at the NTCIR-16 FinNum-3 Task: **Data Augmentation for Financial Numclaim Classification**











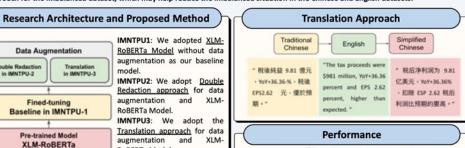


Yung-Wei Teng 1, Pei-Tz Chiu 1, Ting-Yun Hsiao 1, Mike Tian-Jian Jiang 2 and Min-Yuh Day 1,* ¹ Information Management, National Taipei University, New Taipei City, Taiwan

² Zeals Co., Ltd. Tokyo, Japan

myday@gm.ntpu.edu.tw

This paper provides a detailed description of IMNTPU team at the NTCIR-16 FinNum-3 shared task in formal financial documents. We proposed the use of the XLM-RoBERTa-based model with two different approaches on data augmentation to perform the binary classification task in FinNum-3. The first run (i.e., IMNTPU-1) is our baseline through the fine-tuning of the XLM-RoBERTa without data augmentation. However, we assume that presenting different data augmentations may improve the task performance because of the imbalance in the dataset. Accordingly, we presented double redaction and translation method on data augmentation in the second (IMNTPU-2) and third (IMNTPU- 3) runs, respectively. The best macro-F1 scores obtained by our team in the Chinese and English datasets are 93.18% and 89.86%, respectively. The major contribution in this study provide a new understanding toward data augmentation approach for the imbalanced dataset, which may help reduce the imbalanced situation in the Chinese and English datasets.



RoBERTa Model. Chinese Dataset **English Dataset** Dev Set **Dev Set** Test Set **Tokenization Tricks** F1-Score F1-Score F1-Score (%) (%) (%) Input: Good day and welcome to the Apple Inc. Third IMNTPU1 90.51 93.18 87.13 Quarter Fiscal Year 2018 Earnings Conference Call. Today's call is being recorded. IMNTPU2 88.65 91.64 88.82 XLM-RoBERTa Tokenizer **IMNTPU3** 92.16 91.64 Output: <s> Good day and Output: <s> <mask> Good welcome to the Apple day and <mask> to the **Conclusions and Contributions** Inc. Third Quarter Apple <mask>

Quarter Fiscal Year xxnum

2018 Earnings Conference Call. Today's call is

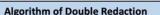
<mask> recorded. </s>

The performance with data augmentation method (Double Redaction) in English dataset is superior than without data augmentation.

- · The major contribution of the research is that data augmentation approach may help reduce imbalanced
- We have developed a novel method for data augmentation technique, which is double redaction and translation approach, and can decrease the issue of imbalanced dataset.

ACKNOWLEDGMENTS

This research was supported in part by the Ministry of Science and Technology (MOST), Taiwan under grant number 110-2410-H-305-013-MY2, and National Taipei University (NTPU) under grant number 110-NTPU-ORDA-F-001 111-NTPU-ORDA-F-001 and 111-NTPU-ORDA-F-003



1: Shuffle the tokens in sentence 2. Delete the duplicated tokens in sentence 3: Copy the remaining tokens as B

Conference

4: SET the δ and γ 5: for specific token in β do

Fiscal Year xxnum 2018

Call. Today's call is

being recorded. </s>

- if γ less than δ then Replace original token with <usk> token
- Cover original token as «mask» token
- end if ii: end for 12: while True do
- Model predict the original token of <usk> and <mask>

Information Management, National Taipei University 《 國本基本大學 2

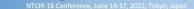
Test Set

F1-Score

(%)

88.39

89.86







NTCIR-17 Best Poster Presentation Award

NTCIR-17 FinArg-1

| IMNTPU at the NTCIR-17 FinArg-1 | Argument-based Sentiment Analysis and Identifying Attack and Support Argumentative Relations in Social Media Discussion Threads

















Chia-Tung Tsai¹, Wen-Hsuan Liao¹, Hsiao-Chuan Liu¹, Vidhya Nataraj², Tzu-Yu Liu³, Mike Tian-Jian Jiang⁴ and Min-Yuh Day¹,*

*myday@gm.ntpu.edu.tw

¹Information Management, ²Smart Healthcare Management, ³Business Administration, National Taipei University, New Taipei City, Taiwan

⁴Zeals Co., Ltd. Tokyo, Japan

In recent years, there has been a surge of interest in argument-based sentiment analysis and the identification of argumentative relationships in social media. These tasks encompass sentiment analysis of premises and claims, as well as the classification of argumentative relationships. Within these tasks, we have developed a fine-tuning method for transformer models. To evaluate and showcase this concept, we established a comprehensive framework to test and display the performance of BERT, ROBERTA, FINBERT, ALBERT, and GPT 3.5-turbo models on financial data and social media texts. Ultimately, the experimental results of these sub-tasks validate the effectiveness of our strategies. The primary contribution of our research is our proposal of two key elements: fine-tuning predominantly with BERT models and employing GPT for generative classification, aiming to enhance the identification of argumentative classifications. Through fine-tuning techniques, the state-of-the-art models can achieve better performance than the baseline.

Transformer-Based Pretrained Model Fine-tuning Techniques BERT-Based Optimization GPT 3.5-turbo Assist Improvement

Fine-tuning Techniques

- Our research in Natural Language Processing (NLP) explores deep learning models like BERT, ALBERT, and ROBERTa for sentence classification. ROBERTa, in particular, shows superior performance in NLP tasks due to more data and extended training, refining BERT's original training approach. The study used RobertaTokenizer for tokenization and RobertaForSequenceClassification for training and evaluation.
- A 5-fold cross-validation technique was employed to fine-tune and assess model performance, involving dividing the dataset into five parts and using each in turn for validation. This ensures a stable and reliable performance evaluation. The study also adjusted hyperparameters such as sentence length, batch size, and training epochs to improve learning efficiency. For fair comparison, the same settings were applied to both ROBERTa and BERT models during fine-tuning.

GPT Generation Strategies and Optimization

- In our study, we demonstrate the application of OpenAl's ChatGPT
 API, integrating deep learning with Natural Language Processing
 (NLP) for detailed text analysis. The technology is finely tuned to
 classify sentences accurately as either "claim" or "premise", aiding
 researchers in identifying core arguments and their supporting
 reasons. This classification is part of a multi-step process, with
 specific sentiment labels providing clear targets for the model.
- The distinction between "claim" and "premise" is vital for understanding arguments and their justifications. Moreover, the method's scalability and adaptability make it versatile, suitable for not only basic sentiment analysis but also for more complex text analysis with additional classification labels

Hyperparameter Settings					
NTCIR-17 FinArg-1 Hyperparameter Settings					
Hyperparameter	Value				
Learning Rate	1e-5, 5e-5				
Max Length	128, 256				
Batch Size	8, 16				
Epochs	3, 4, 5				

	P	erforma	nce		
NT	CIR-17 FinArg	g-1 Argument	Unit Classification	1	
Model	Micro-F1	Macro-F1	Weight-F1	Accuracy	
IMNTPU-1 (BERT-base)	75.44%	75.31%	75.40%	74.82%	
IMNTPU-2 (RoBERTa-base)	76.06%	76.05%	76.07%	75.64%	
IMNTPU-3 (GPT 3.5-turbo)	56.97%	56.82%	56.70%	55.08%	
NTCIR-17 FinArg-1 Argument Relation Detection and Classification					
Model	Micro-F1	Macro-F1	Weight-F1	Accuracy	
IMNTPU-1 (RoBERTa-base)	78.99%	47.36%	76.54%	78.55%	
IMNTPU-2 (FinBERT)	82.61%	52.97%	82.14%	79.13%	
IMNTPU-3 (BERT-uncased)	80.72%	50.73%	79.67%	78.55%	
NTCIR-17 FinArg-1		Attack and Su		tive Relations in	
Model	Micr	o-F1	Macro-F1	Weight-F1	
IMNTPU-1 (Finetuned-Albert)	52.8	38%	34.77%	48.73%	
IMNTPU-2	49.7	71%	24 64%	40 50%	

Conclusions and Contributions

- We combined fine-tuning BERT and RoBERTa with the innovative use of GPT 3.5 Turbo, effectively capturing subtle nuances in conversational texts while demonstrating significant performance in generative tasks.
- Our study offers a comprehensive solution to the Argument Unit Classification challenge, thoroughly evaluating various methods' pros and cons. Additionally, in the multi-class classification task of financial sentiment analysis, we've revealed deeper semantic aspects of texts by analyzing inter-sentential relationship.

ACKNOWLEDGMENTS

(RoBERTa-Large)

This research was supported in part by the National Science and Technology Council (NSTC), Taiwan, under grants MOST 110-2410-H-305-013-MY2, NSTC 112-2425-H-305-002, and NSTC 112-2627-M-038-001, and National Taipel University (NTPU), Taiwan under grants 112-NTPU-ORDAF-03. 112-NTPU-ORDAF-040. USTP-NTPU-TIM-112-101. NTPU-112A113610. and NTPU-112A113613.











NTCIR-17 **Best Poster**

Presentation

Award

NTCIR-17 Real **MedNLP**

IMNTPU at the NTCIR-17 Real-MedNLP Task: Multi-Model Approach to Adverse Drug Event Detection from **Social Media**

















Hsiao-Chuan Liu¹, Vidhya Nataraj², Chia-Tung Tsai¹, Wen-Hsuan Liao¹, Tzu-Yu Liu³, Mike Tian-Jian⁴ Jiang and Min-Yuh Day^{1, 4} *mvdav@gm.ntpu.edu.tw

> ¹Information Management, ²Smart Healthcare Management, ³Business Administration, National Taipei University, New Taipei City, Taiwan ⁴Zeals Co., Ltd. Tokyo, Japan

The IMNTPU team engaged in the NTCIR-17 RealMedNLP task, specifically focusing on Subtask1: Adverse Drug Event detection (ADE) and the challenge of identifying related radiology reports. This task is centered on harnessing methodologies that offer significant aid in real-world medical services, especially when training resources are limited. In our approach, we harnessed the power of pre-trained language models (PLMs), particularly leveraging models like the BERT transformer, to understand both sentence and document structures. Our experimentation with diverse network designs based on PLMs paved the way for an enlightening comparative analysis. Notably, BioBERT-Base emerged as a superior contender, showcasing commendable accuracy relative to its peers. Furthermore, our investigation made strides in the realm of oneshot learning for multiclass labeling, specifically with the GPT framework. The insights gathered emphasized the necessity for more specialized strategies, suggesting avenues for future research in multiclass labeling tasks.

Research Architecture Fine-tuning Techniques RoBERTa-base Best Performani Models MedNLP-SC-SM-EN Evaluation Metrics Accuracy Precision Recall F1 Score Large Language 1-shot Prompt

Prompt Engineer Fine-tuning Techniques

One-shot Learning Analysis showed reduced

accuracy in insight extraction from short, ambiguous tweets. GPT models often over-labeled:

GPT-3.5 labeled 929 instances, GPT-4.0 labeled 789, while the actual ground truth was 400.

Hyperparameters Fine-tuned for multi-label text

classification

Max Epochs: 10

Max Sequence Length: 512 Learning Rate: 5e-5

Batch Size: 16 Loss Func.: BCEWithLogitsLoss

You are a medical expert analyzing tweets to check whether the You are a meanch expert analyzing tweets to check whether the "Scenario" security of the security security and security are short. Please consider this situation and annotate the text with proper labels to check whether the user suffers adverse drug events. For instance, users list the adverse drug events.

reactions.
**Your annotating steps are as follows: **
1. Check whether the user lists the adverse drug effects rather than expressing personal experiences of adverse . Check whether this tweet's user suffers from adverse drug

events.

3. Check the symptoms in these 22 symptoms listed below

**Your annotation should be in the following format: ** If the user suffers from the tweet's symptom instead of listing the adverse drug events, output with the corresponding

label. 2. If the user doesn't suffer from the symptom in the tweet, output with \"None\".

Symptom Labels:

"Symptom Raders." rash, stomatitis mauses, diarrhea." rash, stomatitis "*Here is some annotate example for you to base on. ** Text: I finished C due to side effects of the contrast dye. I was feeling kind of sick and nausea was getting worse, so I thught it woul be tough, but this morning my chest hurts... I thught it would be tough, but this morning my chest hurts... I shall a nauseau so nors left until the test results... See all the seal to such that the seal to such the seal to such that the seal that the seal to such that the seal Label: nausea, pain es for GPT to know.

Performance

Exact Accuracy of Test Dataset and Development Dataset							
1	Models	Dev	Development Dataset (#1,192) Test Dataset		(#1,993)		
	BERT-Base ission Run 1)		0.92			0.82	
Rob	erta-Base		0.76			-	
	erta-Large ission Run 2)		0.85			0.81	
	SPT3.5 ission Run 3)		0.72 0.69				
(SPT 4.0		0.62 -				
Subtask 1-SM-ADE-EN Binary and Per Label Performance Metrics							
Models Score .	Pred	Precision Rec		all F1 Score		core	
wodels	Score	ADE	NO ADE	ADE	NO ADE	ADE	NO ADE
BioBERT-	Binary	0.74	0.91	0.78	0.89	0.76	0.90
Base	Per label	0.72	1.00	0.76	0.99	0.74	0.99
RoBERTa-	Binary	0.73	0.93	0.83	0.88	0.78	0.90
	Per label	0.71	1.00	0.77	0.99	0.74	0.99
Large	rei iabei						

0.47 0.47 0.20 **0.91** GPT3.5 0.42 0.98 0.18 **1.00** Subtask 1-SM-ADE-EN Binary and Per Label Performance Metrics in

Development Dataset ADE NO ADE ADE NO ADE ADE NO ADE BioBERT-0.90 0.97 0.93 0.96 0.92 0.97 Per label 0.97 0.93 0.96 0.92 Roberts. 0.89 0.83 0.59 0.97 0.71 Base Per label 0.85 0.99 0.49 1.00 0.62 0.85 0.94 0.86 RoBERTa-Binary 0.87 0.93 0.81 Per lahel 0.84 0.99 0.77 1.00 Binary 0.72 0.80 0.53 0.90 0.61 0.62 0.99 0.57 0.99 0.60 0.52 0.96 0.95 0.61 0.67 0.75 0.48 **1.00** 0.92 0.98 0.63

Exact Match Accuracy Results in Development Dataset						
Models	Accuracy	Models	Accuracy			
RoBERTa-Base	0.86	GPT3.5-1-shot	0.54			
RoBERTa-Large	0.87	GPT3.5-scenario	0.69			
BioBERT-Base	0.85	GPT4.0-1-shot	0.61			
BioBERT-Large	0.86	GPT4.0-scenario	0.70			
		GPT3.5-fintune	0.85			

Conclusions and Contributions

- Expanded dataset with GPT-3.5 boosts RoBERTa accuracy from 0.76 to 0.86 Refined prompts for tweets, increasing GPT4.0 accuracy to 0.70.
- BioBERT excels in drug event extraction; GPT one-shot learning shows limits. We proposed two critical elements, Prompt Engineer and Fine-Tuning Technique GPT-4.0 showing enhanced performance in ambiguous datasets.

ACKNOWLEDGMENT

This research was supported in part by the National Science and Technology Council (NSTC). Taiwan, under grants MOST 110-2410-H-305-013-MY2, NSTC 112- 2425-H-305-002-, and NSTC 112-2627-M-038-001-, and National Taipei University (NTPU), Taiwan under grants 112-NTPU-ORDA-F-003. 112- NTPU-ORDA-F-004. USTP-NTPU-TMU-112-01. NTPU-112A413E01. and NTPU-112A513E01









2024

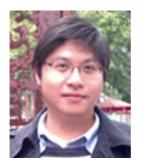


Generative AI in Multimodal Cross-Lingual Dialogue System for Inclusive Communication Support















Vidhya Nataraj

Wen-Hsuan Liao

Yue-Shan Chang

Chen-Yu Chiang

Chao-Yin Lin

Yu-An Lin

Min-Yuh Day*

Institute of Information Management, National Taipei University, Taiwan

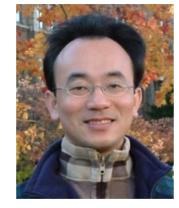
^{*} myday@gm.ntpu.edu.tw



2024



Smart City Large Language Model Agent System



Min-Yuh Day *



Xin-Ting Lu



Xu-Yu Lan



Bor-Jen Chen

* myday@gm.ntpu.edu.tw

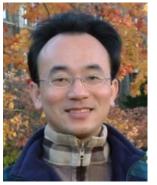
Institute of Information Management,
National Taipei University, Taiwan



Fintech and Green Finance Center (FGFC), NTPU



Fintech Green Finance for Carbon Market Index, Corporate Finance, and Environmental Policies (金融科技綠色金融於碳市場指標與公司環境策略)



戴敏育 資管所教授 研究中心主任



額汝芳 統計系副教授 研究中心副主任



游擱嘉 企管系助理教授 研究中心執行長



翁新傑 統計系助理教授 研究中心課程組組長



黃懷陞 資管所助理教授 研究中心產學組組長

Summary



- Overview of ESG
 - Environmental, Social, Governance
- Challenges in ESG Data Analytics
 - Fragmented, Unstructured, Diverse
- Generative AI as a Solution
 - Synthesis, Reporting, and Prediction
- Sustainability Innovation



Acknowledgments: Research Projects



- 1. Fintech Green Finance for Carbon Market Index, Corporate Finance, and Environmental Policies.

 Carbon Emission Sentiment Index with AI Text Analytics
 - NTPU, 113-NTPU_ORDA-F-003, 2023/01/01~2024/12/31
- 2. Digital Support, Unimpeded Communication: The Development, Support and Promotion of AI-assisted Communication Assistive Devices for Speech Impairment (2/3).
 - Multimodal Cross-lingual Task-Oriented Dialogue System for Inclusive Communication Support
 - NSTC 113-2425-H-305-002-, 3 Years (2023/05/01-2026/04/30) Year 1: 2024/05/01~2025/04/30
- 3. Research on speech processing, synthesis, recognition, and sentence construction of people with language disabilities. Multimodal Cross-lingual Task-Oriented Dialogue System
 - NTPU, 113-NTPU_ORDA-F-004, 2023/01/01~2025/12/31
- 4. Metaverse AI Multimodal Cross-Language Task-Oriented Dialogue System
 - ATEC Group, Fintech and Green Finance Center (FGFC, NTPU), NTPU-112A413E01, 3 Years (2023/05/01~2026/04/30)
- 5. Generative Al-Driven ESG Report Generation Technology
 - Industrial Technology Research Institute (ITRI), Fintech and Green Finance Center (FGFC, NTPU), NTPU-113A513E01, 2024/03/01~2024/12/31
- 6. Establishment and Implement of Smart Assistive Technology for Dementia Care and Its Socio-Economic Impacts (3/3). Intelligent, individualized and precise care with smart AT and system integration
 - NSTC, 113-2627-M-038-001-, 2024/08/01~2025/07/31
- 7. Prospective longitudinal study on peri-implant bone loss associated with peri-implantitis
 - USTP (NTPU, TMU), USTP-NTPU-TMU-113-03, 2024/01/01~2024/12/31



Acknowledgments: IFIT Lab Members





Intelligent Financial Innovation Technology, IFIT Lab, IM, NTPU

References



- Cino Robin Castelli, Cyril Shmatov (2022), Quantitative Methods for ESG Finance, Wiley
- Simon Thompson (2023), Green and Sustainable Finance: Principles and Practice in Banking, Investment and Insurance, 2nd Edition, Kogan Page.
- Chrissa Pagitsas (2023), Chief Sustainability Officers At Work: How CSOs Build Successful Sustainability and ESG Strategies, Apress.
- Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun (2023). "A Comprehensive Survey of Al-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT." arXiv preprint arXiv:2303.04226.
- Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. (2023) "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing." ACM Computing Surveys 55, no. 9 (2023): 1-35.
- Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min et al. (2023) "A Survey of Large Language Models." arXiv preprint arXiv:2303.18223.
- Touvron, Hugo, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov et al. (2023) "Llama 2: Open Foundation and Fine-Tuned Chat Models." arXiv preprint arXiv:2307.09288 (2023).
- Rafailov, R., Sharma, A., Mitchell, E., Ermon, S., Manning, C. D., & Finn, C. (2023). Direct preference optimization: Your language model is secretly a reward model. arXiv preprint arXiv:2305.18290.
- Tunstall, Lewis, Edward Beeching, Nathan Lambert, Nazneen Rajani, Kashif Rasul, Younes Belkada, Shengyi Huang et al. "Zephyr: Direct Distillation of LM Alignment." arXiv preprint arXiv:2310.16944 (2023).
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. arXiv preprint arXiv:2203.02155.
- Gozalo-Brizuela, Roberto, and Eduardo C. Garrido-Merchan (2023). "ChatGPT is not all you need. A State of the Art Review of large Generative AI models." arXiv preprint arXiv:2301.04655 (2023).
- Wenliang Dai, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Junqi Zhao, Weisheng Wang, Boyang Li, Pascale Fung, and Steven Hoi. (2023) "InstructBLIP: Towards General-purpose Vision-Language Models with Instruction Tuning." arXiv preprint arXiv:2305.06500 (2023).
- Shahab Saquib Sohail, Faiza Farhat, Yassine Himeur, Mohammad Nadeem, Dag Øivind Madsen, Yashbir Singh, Shadi Atalla, and Wathiq Mansoor (2023). "The Future of GPT: A Taxonomy of Existing ChatGPT Research, Current Challenges, and Possible Future Directions." Current Challenges, and Possible Future Directions (April 8, 2023) (2023).
- Longbing Cao (2022). "Decentralized ai: Edge intelligence and smart blockchain, metaverse, web3, and desci." IEEE Intelligent Systems 37, no. 3: 6-19.
- Qinglin Yang, Yetong Zhao, Huawei Huang, Zehui Xiong, Jiawen Kang, and Zibin Zheng (2022). "Fusing blockchain and AI with metaverse: A survey." IEEE Open Journal of the Computer Society 3: 122-136.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. arXiv preprint arXiv:2203.02155.



Tsukuba x NTPU

Bilateral Academic Exchange Program

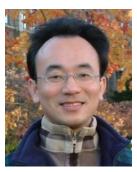




Q & A

Generative AI for ESG Data Analytics and Sustainability Innovation





Min-Yuh Day, Ph.D.

Professor, Institute of information Management

Director, Fintech and Green Finance Center (FGFC)

Division Director, Sustainable Development, Sustainability Office

National Taipei University

https://web.ntpu.edu.tw/~myday

