

黃東浩博士

萬事達卡研發副總裁兼全球新興技術負責人 | AI 領域高管學者

donghao.huang@gmail.com | 美國弗吉尼亞州阿靈頓

新加坡公民 | 專利：[10 項已授權](#)，[50+ 項審批中](#)

職業概述

資深科技高管，擁有 27 年以上從業經驗，橫跨軟體工程、Web3、AI/ML/生成式 AI 與支付創新領域。融合頂尖企業實戰與前沿學術研究，近期於新加坡管理大學完成人工智慧工程博士學位。發表 28 篇同行評審論文，發明 60 餘項專利，專注於將前沿 AI 突破轉化為可規模化的金融基礎設施。

教育背景

人工智慧工程博士 (Doctor of Engineering)

2023 年 8 月 – 2026 年 5 月

新加坡管理大學，新加坡

- 博士論文：[Generative AI in Enterprises: Optimizing Applications with Large Language Models](#) (244 頁)。主導師：Prof. Zhaoxia Wang; 副導師: Prof. Chong Wah Ngo.
- 2025 年獲新加坡資訊通訊媒體發展局 (IMDA) 數字獎學金 (研究生)
- 專題報導：<https://computing.smu.edu.sg/engd/community-stories/engineering-future-ai-smu-engd-powers-innovation-global-technology-leader>
- 研究方向：檢索增強生成 (RAG)、邊緣 AI、LLM 和 ML 模型最佳化、多智慧體系統、情感分析、推理型 LLM
- 在 AI 學術會議和期刊發表 25 篇同行評審論文 (AAAI、AAMAS、NAACL、IJCNN、PAKDD、IEEE CAI、IEEE SSCI、IEEE ICDM、IEEE Intelligent Systems、Artificial Intelligence Reviews、ACM Computing Surveys)

軟體工程技術碩士

2004 年 1 月 – 2007 年 11 月

新加坡國立大學，新加坡

新加坡經濟發展局(EDB)研究與培訓專案證書

2001 年 9 月 – 2002 年 8 月

新加坡國立大學光電子中心，新加坡

理論物理學碩士

1998 年 9 月 – 2001 年 7 月

北京大學，中國北京

- 在物理學術期刊發表 2 篇同行評審論文

工程物理學學士

1989 年 9 月 – 1995 年 7 月

清華大學，中國北京

職業經歷

研發副總裁 / 全球新興技術負責人 – Mastercard (美國)

2025 年 8 月 – 至今

美國弗吉尼亞州阿靈頓

- 領導 AI 新興技術 (EmTech) 領域的全球研發計劃
- 領導 5 個戰略性"大賭注"專案的 EmTech 研究——旨在塑造 Mastercard 下一代支付基礎設施和 AI 驅動商業解決方案的高影響力機密專案
- 對話式支付智慧體 (CPA / ConvPayMAS) ——構建並領導智慧體商務支付技術棧，融合 Google A2A 智慧體間協議、AP2 三重授權鏈 (Intent → Cart → Payment) 與 MCP 工具服務，採用分層 LangGraph 主管智慧體調度卡片註冊、選卡、驗證與扣款工作流，並由 PCI/PII 防護智慧體保障合規。發表 3 篇論文 (PAKDD 2026 主會、PAKDD 2026 AIDS4DF 研討會、AAMAS 2026 演示論文)，提交 3 項專利申請。
 - 提出智慧體成功率 (ASR) 指標——一種比較實際與預期智慧體轉移序列的軌跡保真度量，而非僅看最終結果；在 9 萬次任務實例 (18 個 LLM × 4 個場景 × 250 個用例 × 5 次重複，約 1.15 億 tokens) 的評測中，發現 18 個模型中有 10 個會靜默跳過支付確認環節——該缺陷在任務成功率與交接指標下完全不可見。
 - 依據上述診斷，將本地部署開源模型的支付任務成功率從 4.0% 提升至 93.4% (Llama 3.1-8B)、從 3.2% 提升至 94.7% (Magistral-24B)，大幅縮小與 GPT-4.1 的差距。
- GAIME (生成式 AI 商戶資訊增強) ——領導面向 Mastercard 交易描述符的商戶資訊抽取與實體匹配研究專案，該生產系統日均處理數百萬筆交易。發表 1 篇論文 (PAKDD 2026 RAFDA 研討會)，1 篇已錄用 (IEEE ICDM 2026)，1 篇在審 (EACL 2027 產業論文)。
 - 在 8,015 條標註交易描述符上對四大模型家族 (Gemma 3 270M-4B、Qwen 3.5 0.8B-4B、Aya 3.35B、Llama 3.1-8B) 的 24 個變體進行 LoRA 微調；以 4B 模型達到 96.60% F1，媲美現有 Llama-3.1-8B 生產基線 (96.95% F1)，參數量減半且有效推理速度提升 1.2 倍，並通過 MLflow 將 14 個變體部署為 Databricks 模型服務端點。
 - 在 8 個實體匹配資料集上對 46 種 LLM 配置進行基準評測——開源 GPT-OSS-120B 以約 1/20 的成本達到 92.1 F1，媲美高推理強度的 GPT-5；並識別出可媲美現有生產判別模型 (F0.5 94.8) 且延遲減半的本地方案。
- 組織並主持 SENTIRE'25 (第 15 屆情感分析與語言關聯資料研討會)，與 IEEE ICDM 2025 聯合舉辦，2025 年 11 月 13 日，美國華盛頓特區

研發副總裁 / 全球新興技術負責人 – Mastercard 亞太

2022 年 5 月 – 2025 年 7 月

新加坡

- 領導新興技術 (EmTech) 領域的全球研發計劃：AI / 機器學習 / 生成式 AI 和 Web3
- AI/LLM/GenAI 專案：
 - 領導 Mastercard 助手平臺 (MAP) 的 EmTech 研究，這是一套旨在簡化 Mastercard 產品客戶入門流程的數字助手

- 使用 LangChain、FAISS 向量資料庫和 HuggingFace 嵌入構建初始 RAG 原型
- 在 NVIDIA A40 和 L40 GPU 上對開源 LLM (Llama 2、Orca 2) 進行早期基準測試，展示 24-49% 的推理速度提升
- 開發創新的重複感知效能 (RAP) 評估框架，現已部署於 MAP 生產環境
- 領導使用 Ollama 在多種硬體平臺上進行邊緣感知 LLM 部署研究
- 使用 LLM (GPT-4/5、Claude、Llama、Qwen) 開創企業 GenAI 應用，實現支付/商業自動化
 - 使用 QLoRA 在多 GPU 配置上微調 70B+ 引數模型，用於中英翻譯和邏輯推理
 - 開發多智慧體系統，使用邊緣部署 LLM 實現 99.9% 準確率的發票對賬
 - 開發 HMASP——首個實現端到端智慧體支付工作流的基於 LLM 的多智慧體系統
 - 開發 CMAS4G2 協作多智慧體系統，用於自動生成 Gherkin 驗收測試
 - 建立 LLM-as-a-Judge 評估框架，相比 GPT-5 實現 47 倍成本降低
- 經典機器學習、預測與風險建模專案：
 - 構建結合經典 ML 模型 (邏輯回歸、SVM、隨機森林、XGBoost) 與工程化及 LLM 衍生特徵的金融決策流水線——採用網格搜尋超引數調優、5 折交叉驗證、序列特徵選擇和對數損失概率校準，驅動風險敏感的凱利準則下注規模
 - 為 ICO / 加密專案成功預測開發整合模型 (樸素貝葉斯、SVM、邏輯回歸、隨機森林、CatBoost、神經網路)，結合基於情感和參與度的特徵工程
 - 對完整建模範圍進行基準測試——傳統 ML (隨機森林、基於 TF-IDF 的 SVM)、深度學習 (CNN、RNN、LSTM) 和 Transformer (BERT、RoBERTa) ——用於假新聞檢測和自動化海事風險分類，引入新穎的輸出可靠性指標：有效類別比率
- Web3 專案：
 - 開發 Mastercard 代幣化卡片 NFT 解決方案，支援將支付卡作為靈魂繫結代幣 (SBT) 進行代幣化
 - 開發 DigiPay，利用 Layer 2 區塊鏈和 ERC-4337 帳戶抽象的 Web3 微支付穩定幣解決方案

研發副總裁 / 新加坡研發負責人 – Mastercard 亞太

2011 年 8 月 – 2022 年 5 月

新加坡

- 建立並領導新加坡研發團隊——Mastercard 全球創新與研發部門；管理位於新加坡的高績效研發團隊
- 從事支付/商業領域的創新與研發，涵蓋前沿技術：EMV 晶片、物聯網、二維碼支付、車聯網、人形機器人、對話式 AI、AR/VR、區塊鏈和數字資產、Web3、NFT、DeFi、元宇宙
- 推動複雜專案從創意到商業化交接的全流程；促進專案向 Mastercard 商業化流程的過渡
- 作為技術創新的倡導者，與各類合作伙伴 (內部部門、外部合作伙伴、大學) 協作，確保 Mastercard 在支付領域建立技術卓越的聲譽
- 主要商業化專案：

- 2021 年 11 月：Mastercard 與新奧爾良市和 MoCaFi 合作推出 Crescent City Card 專案
- 2020 年 9 月：Mastercard 與三星、Airtel 和 Asante 合作，透過按需付費服務推動非洲數字普惠
- 2018 年 10 月：Mastercard、Centenary Bank 和 M-KOPA Solar 在烏幹達推出首創能源解決方案
- 2018 年 2 月：Mastercard 使用 Facebook Messenger 幫助小企業數字化
- 2017 年 9 月：SPH Buzz 與 Mastercard 合作推出新型混合商店概念
- 2017 年 2 月：BharatQR——全球首個可互操作的二維碼受理解決方案在印度推出
- 2016 年 5 月：MasterCard 為軟銀機器人形機器人"Pepper"提供首個商業應用
- 2015 年 9 月：MasterCard 變革援助物資分發方式
- 2012 年 8 月：StarHub 與 DBS、EZ-Link 和 MasterCard 合作推出智慧錢包

高階軟體工程師 – PayPal

2008 年 7 月 – 2011 年 8 月

新加坡

- 領導軟體開發團隊，負責一系列針對巴西市場的專案，以滿足這一快速增長市場的需求
- 限額/驗證/合規領域的主題專家
- 主要商業化專案：
 - 領導 12 人開發團隊，在 8 個月內交付最大的歐盟合規專案 (EU KYB Receiving Limit)
 - 構建中國銀聯 (CUP) 整合，實現實時交易處理

軟體研發經理 – Streaming21 (新加坡)

2007 年 10 月 – 2008 年 7 月

新加坡

- 建立研發中心並領導媒體中繼伺服器的開發——一款領先的 IPTV 流媒體伺服器

軟體開發負責人 – Schmidt Electronics (S.E.A.)

2006 年 1 月 – 2007 年 9 月

新加坡

軟體工程師 – Unaxis 新加坡

2005 年 2 月 – 2005 年 12 月

新加坡

軟體工程師 – ASM Technology 新加坡

2002 年 9 月 – 2005 年 2 月

新加坡

軟體工程師 – 北京智達科技

1995 年 8 月 – 1998 年 8 月

中國北京

演講與論壇

- 受邀演講，IMDA AI Agents 技術分享會，新加坡 (2026 年 4 月 23 日) —— 與南洋理工大學校長講席教授、AAAI Fellow 安波教授同台

- 受邀演講，InnoAI 2026 全球人工智慧創新未來峰會 —— 與摩根大通、谷歌、輝達、eBay 等公司嘉賓同列
- 研討會主席，SENTIRE' 25 (第 15 屆情感分析與語言鏈結資料研討會)，與 IEEE ICDM 2025 聯合舉辦，2025 年 11 月，美國華盛頓特區
- 受邀演講，SuperAI Singapore 2025，新加坡濱海灣金沙 (2025 年 6 月 19 日) —— “企業級開源大語言模型優化與基準測試”
- 受邀論壇嘉賓，IEEE 人工智慧大會 (IEEE CAI 2024)，新加坡濱海灣金沙 (2024 年 6 月) —— 論壇由 Simon See (輝達 AI 技術中心總監) 主持，同台嘉賓包括杉山將教授 (日本理化學研究所先進智能研究中心主任)。800+全球參會者
- 獨立演講及論壇嘉賓，CALM SUMMIT' 24 (Rasa)，2024 年 10 月 —— “面向對話式 AI 的開源大語言模型優化與基準測試” + 論壇 “可擴展語言模型的優勢”
- 主旨演講，Token 2049 — WSPN Stable Rise，新加坡 (2024 年 9 月) — “構建信任與流暢商業體驗：引領 Web3 技術支付未來”
- 受邀演講，FCon 2023 全球金融科技大會，中國上海 (2023 年) — 代表萬事達卡實驗室 — “Web3 時代的商業未來”。影片：<https://www.infoq.cn/video/3UlpCdtmzxz5OwRjd6j9>
- 受邀演講，Upbit 開發者大會 (UDC) 2022，韓國釜山 (2022 年 9 月) —— “萬事達卡為 Web3 和元宇宙做好準備”，代表萬事達卡出席，同場嘉賓包括孫宇晨 (波場創始人)、Sebastien Borget (The Sandbox COO)、Bryan Pellegrino (LayerZero Labs CEO) 及 Leemon Baird (Hedera 聯合 CEO)
- 論文發表演講：PAKDD 2024/2025/2026 (台北/雪梨/香港)、ICDAAI 2026 (2026 年 5 月，中國成都)、AAMAS 2026 (2026 年 5 月，賽普勒斯帕福斯)、IEEE ICDM 2025 (2025 年 11 月，美國華盛頓特區)、IJCNN 2025 (2025 年 6 月，義大利羅馬)、NAACL 2025 (2025 年 4 月，美國阿爾伯克基)、IEEE SSCI 2025 (2025 年 3 月，挪威特隆赫姆)、IEEE CAI 2024 (2024 年 6 月，新加坡濱海灣金沙)

榮譽與獎項

- IMDA 新加坡數位獎學金 (研究生)，2025 —— 新加坡政府頒發的競爭性獎學金，授予從事高級研究的資深技術專業人員
- 新加坡管理大學圖書館 "Generate Your L(AI)brary" 黑客松一等獎，2023 —— 在新加坡管理大學 AI 黑客松中獲得第一名 (2023 年 8 月 26-27 日)：https://www.linkedin.com/posts/donghao-huang_yesterday-our-team-won-1st-prize-in-smu-library-ugcPost-7101786441805926400-LW0S/
- Smart Sentosa 挑戰賽冠軍，2016 —— 萬事達卡亞太區智慧城市引擎贏得 Smart Sentosa 挑戰賽，該賽事為新加坡政府智慧城市創新競賽

精選媒體報導

- “Engineering the Future of AI: SMU EngD Powers Innovation for a Global Technology Leader” — 新加坡管理大學計算機與資訊系統學院，2026 年 2 月。

<https://computing.smu.edu.sg/engd/community-stories/engineering-future-ai-smu-engd-powers-innovation-global-technology-leader>

- Token 2049 Side Event Coverage — Investing.com · 2024 年 9 月。
<https://hk.investing.com/analysis/article-110563>
- “支付科技與安全交易” 論壇 — 新浪財經 · 2018 年 3 月。
<https://finance.sina.cn/blockchain/2018-04-02/detail-ifysvnxn1920535.d.html>
- “零距離感受萬事達卡的支付黑科技” — 新浪/支付產業網 · 2017 年 11 月。
https://k.sina.cn/article_6007803312_16617cdb0001001wwd.html
- “不換手機和掏不起錢的人，怎樣才能用上新支付技術？” — 中國電子銀行網/cnBeta · 2017 年 11 月。
<https://www.cebnet.com.cn/20171127/102444564.html>
- “亞洲最大！萬事達卡支付科技創新，9 間實驗室撐腰” — Business Next (數位時代) · 2017 年 11 月。
<https://www.bnext.com.tw/article/47137/mastercard-innovation-lab>

志願服務經歷

會長 – 寶嚴 (新加坡) 教育協會	2025 年 11 月 – 至今
副會長 – 北京大學新加坡校友會	2016 年 6 月 – 2020 年 6 月
導師 – Startupbootcamp 金融科技新加坡	2015 年 1 月 – 2019 年 1 月
IT 與支付顧問 – ConneXions 國際 (新加坡)	2011 年 3 月 – 2015 年 12 月
技術顧問 – 聖安德烈自閉症中心 (新加坡)	2006 年 11 月 – 2008 年 7 月

技術技能

- AI/ML/GenAI：LLM (GPT-4/5、Claude、Gemini、Llama、Qwen、DeepSeek-R1、Mistral、Magistral、Phi、Nemotron)、PyTorch、Transformers、PEFT、LoRA/QLoRA 微調、LlamaFactory、Unsloth、RAG、LangChain、LangGraph、Autogen、Ollama、LM Studio、FAISS、向量資料庫、少樣本學習、微調、提示工程、情感分析、NLP
- GPU/硬體：NVIDIA H100、L40、A40、RTX A6000、RTX 4080、RTX 4090、Jetson AGX Orin、DGX Spark；Apple M 系列；4 位量化、模型最佳化、邊緣 AI 部署
- 程式語言與框架：Python、C/C++、Java、NodeJS、React、Vue、Spring Boot/Cloud、Solidity
- 雲與基礎設施：AWS、Azure、Heroku；微服務、CI/CD、Docker
- Web3/區塊鏈：區塊鏈、IPFS、NFT、DeFi、x402 協議

專利

(10 項已授權，50+ 項待審 — 均歸屬萬事達卡)

已授權專利

- **US12417452B2** – Smart chip payment acceptance (2025)
- **US12217246B2** – Method and system for use of an EMV card in a multi-signature wallet for cryptocurrency transactions (2025)
- **US11715100B2** – Electronic system and computerized method for verification of transacting parties to process transactions (2023)
- **US11615406B2** – Method and system for providing a service at a self-service machine (2023)
- **US10535034B2** – Item delivery management systems and methods (2020)
- **US10482499B2** – Method for conducting a transaction (2019)
- **US10423949B2** – Vending machine transactions (2019)
- **US10083427B2** – Method for receiving an electronic receipt of an electronic payment transaction into a mobile device (2018)
- **US9836729B2** – Method and system for conducting a payment transaction and corresponding devices (2017)
- **US8712914B2** – Method and system for facilitating micropayments in a financial transaction system (2014)

部分待審申請

- **WO2026063867A1** – Method and system for seamless cross-network physical-digital (phy-gital) experience (2024)
- **US20260080407A1** – Method and system for seamless cross-network physical-digital (phy-gital) experience (2024)
- **WO2024058723A1** – Digitization of payment cards for web 3.0 and metaverse transactions (2022)
- **US20230059546A1** – Access control system (2021)
- **EP4356332A2** – Method and system for mediated cross ledger stable coin atomic swaps using hashlocks (2021)
- **US20210406887A1** – Method and system for merchant acceptance of cryptocurrency via payment rails (2020)
- **US20210158316A1** – Electronic system and computerized method for controlling operation of service devices (2019)
- **US20200175489A1** – Method and system for QR code originated vending (2017)
- **WO2017200643A1** – Method and system for allocating and transacting with multiple non-financial commodities (2016)
- **US20180025332A1** – Transaction facilitation (2016)

... 以及 40+ 項其他待審專利申請，涵蓋支付技術、區塊鏈、AI 和物聯網領域。

精選論文

1. Huang, D., 等. (2026). [模型能有多小？面向金融交易商戶資訊抽取的 270M–8B 模型 LoRA 微調](#). IEEE ICDM 2026.
2. Teo, N., 等. (2026). [大語言模型常識推理綜述](#). ACM Computing Surveys.
3. Chua, J., 等. (2026). [一種新型基於大語言模型的層級多智慧體支付系統](#). PAKDD 2026.
4. Huang, D., 等. (2026). [任務複雜度的重要性：大語言模型情感分析推理的實證研究](#). PAKDD 2026.
5. Huang, D., & Wang, Z. (2025). [使用 DeepSeek-R1 的可解釋情感分析](#). IEEE Intelligent Systems.
6. Huang, D., 等. (2025). [邊緣 LLM：效能與效率評估](#). IJCNN 2025.
7. Huang, D., 等. (2025). [RAP：平衡重複與效能的指標](#). NAACL 2025.
8. Wang, Z., Huang, D., 等. (2025). [中文情感分析綜述：主題、方法與趨勢](#). Artificial Intelligence Review.
9. Lin, W. B., Huang, D. H., 等. (2001). [弱相互作用大質量粒子的非熱產生與亞星系結構](#). Physical Review Letters.

... 及其它 19 篇同行評審論文 (人工智慧領域共 26 篇，物理領域共 2 篇)

DR. DONGHAO HUANG

VP of R&D & Global Emerging Technology Lead at Mastercard | Executive-Scholar in AI

donghao.huang@gmail.com | Arlington, VA, USA

Singapore Citizen | Patents: 10 granted, 50+ pending

PROFESSIONAL SUMMARY

Senior technology executive with 27+ years of experience across software engineering, Web3, AI/ML/GenAI, and payments innovation. Combines elite enterprise execution with cutting-edge academic research, having recently completed a Doctor of Engineering in AI at Singapore Management University. Author of 28 peer-reviewed papers and inventor of 60+ patents, specializing in translating advanced AI breakthroughs into scalable financial infrastructure.

EDUCATION

Aug 2023 – May 2026 Singapore Management University Singapore

Doctor of Engineering (EngD) in AI

- Doctoral dissertation: [Generative AI in Enterprises: Optimizing Applications with Large Language Models](#) (244 pp.). Supervisor: Prof. Zhaoxia Wang. Co-supervisor: Prof. Chong Wah Ngo.
- Awarded IMDA SG Digital Scholarship (Postgraduate) 2025
- Featured story: <https://computing.smu.edu.sg/engd/community-stories/engineering-future-ai-smu-engd-powers-innovation-global-technology-leader>
- Research focus: Retrieval-Augmented Generation (RAG), Edge AI, LLM and ML model optimization, multi-agent systems, sentiment analysis, reasoning LLMs
- Published 25 peer-reviewed papers at top AI conferences and journals (AAAI, AAMAS, NAACL, IJCNN, PAKDD, IEEE CAI, IEEE SSCI, IEEE ICDM, IEEE Intelligent Systems, Artificial Intelligence Reviews, ACM Computing Surveys)

Jan 2004 – Nov 2007 National University of Singapore Singapore

Master of Technology in Software Engineering

Sep 2001 – Aug 2002 Centre for Optoelectronics, NUS Singapore

Certificate, EDB Research and Training Programme

Sep 1998 – Jul 2001 Peking University Beijing, China

Master of Science in Theoretical Physics

- Published 2 peer-reviewed papers in physics journals

Sep 1989 – Jul 1995 Tsinghua University Beijing, China

Bachelor of Engineering in Engineering Physics

PROFESSIONAL EXPERIENCE

Aug 2025 – to date Mastercard (U.S.) Arlington, VA, USA

VP, R&D / Global EmTech Lead

- Lead global R&D initiatives in AI Emerging Technology (EmTech) domains

- Lead EmTech research for 5 strategic "Big Bets" projects – high-impact, confidential initiatives aimed at shaping Mastercard's next-generation payment infrastructure and AI-powered commerce solutions
- Conversational Payment Agent (CPA / ConvPayMAS) – built and led an agentic-commerce payment stack combining Google's A2A inter-agent protocol, the AP2 three-mandate chain (Intent → Cart → Payment) and MCP tool servers, with a hierarchical LangGraph supervisor routing card registration, selection, validation and charge workflows behind a PCI/PII guardrail agent. 3 papers published – PAKDD 2026 main conference, PAKDD 2026 AIDS4DF workshop, and AAMAS 2026 Demonstration Track; 3 patent applications filed.
 - Introduced the Agentic Success Rate (ASR) metric – a trajectory-fidelity measure comparing observed and expected agent transitions rather than final outcomes alone; a 90,000-instance evaluation (18 LLMs × 4 scenarios × 250 cases × 5 repeats, ~115M tokens) exposed 10 of 18 models silently skipping the payment confirmation checkpoint – a failure invisible to task-success and handoff metrics.
 - Used those diagnostics to lift payment task success on locally hosted open-weight models from 4.0% to 93.4% (Llama 3.1-8B) and 3.2% to 94.7% (Magistral-24B), closing most of the gap to GPT-4.1.
- GAIME (GenAI Merchant Enrichment) – lead the research program behind merchant information extraction and entity matching over Mastercard transaction descriptors, a production system processing millions of transactions daily. 1 paper published (PAKDD 2026 RAFDA workshop); 1 accepted (IEEE ICDM 2026); 1 under review (EACL 2027 Industry Track).
 - LoRA fine-tuned 24 model variants across four families (Gemma 3 270M–4B, Qwen 3.5 0.8B–4B, Aya 3.35B, Llama 3.1-8B) on 8,015 labeled transaction descriptors; matched the incumbent Llama-3.1-8B production baseline (96.95% F1) with a 4B model at 96.60% F1 — at half the parameter count and 1.2× faster effective inference, and deployed 14 variants as Databricks Model Serving endpoints via MLflow.
 - Benchmarked 46 LLM configurations across 8 entity-matching datasets – open-weight GPT-OSS-120B reaches 92.1 F1 at ~20× lower cost than GPT-5 at high reasoning effort – and identified local judge configurations matching the deployed production judge (F0.5 94.8) at half the latency.
- Organized and hosted SENTIRE'25 (15th Workshop on Sentiment Analysis and Linguistic Linked Data) co-located with IEEE ICDM 2025, November 13th, Washington DC.

May 2022 – Jul 2025 Mastercard Asia/Pacific Pte Ltd

Singapore

VP, R&D / Global EmTech Lead

- Lead global R&D initiatives in Emerging Technology (EmTech) domains: AI / Machine Learning / Generative AI and Web3.
- **AI/LLM/GenAI Initiatives:**
 - Led EmTech research for [Mastercard Assistant Platform](#) (MAP), a suite of digital assistants designed to simplify the customer onboarding process for Mastercard products
 - Built the initial Retrieval-Augmented Generation (RAG) prototype using LangChain, FAISS vector databases, and HuggingFace embeddings to power enterprise-grade knowledge systems.
 - Conducted early benchmarking of open-source LLMs (Llama 2, Orca 2) on NVIDIA A40 and L40 GPUs (48GB each), demonstrating 24-49% inference speed improvement on L40 vs A40.

- Developed a novel Repetition-aware Performance (RAP) evaluation framework for LLM quality assessment, now fully deployed in the MAP production environment.
- Led research on edge-aware LLM deployment using the Ollama framework across diverse hardware platforms (NVIDIA RTX A6000, NVIDIA RTX 4090, NVIDIA Jetson AGX Orin, Apple M-series—directly informing MAP's deployment strategy.
- Pioneered enterprise GenAI applications using LLMs (GPT-4/5, Claude, Llama, Qwen) for payments/commerce automation
 - Fine-tuned 70B+ parameter models (Llama-3.1-70B, Qwen2-72B) using QLoRA on multi-GPU setups (4× NVIDIA L40 GPUs) for Chinese-English translation and logical reasoning
 - Developed multi-agent systems for invoice reconciliation achieving 99.9% accuracy using edge-deployed LLMs (qwen2.5:32b)
 - Developed hierarchy multi-agent systems for conversational payment agent (HMASP) - first LLM-based multi-agent system to implement end-to-end agentic payment workflows, enabling external AI agents to converse and complete payments using natural language with minimal integration overhead
 - Developed CMAS4G2, a cooperative multi-agent system for automated Gherkin acceptance test generation using Microsoft's AutoGen Swarm framework; demonstrated open-weight models on consumer hardware (NVIDIA RTX 4090, 24GB VRAM) approach proprietary performance (GPT-OSS 20B: 87.6% success vs Claude Sonnet 4: 90.4%)
 - Created LLM-as-a-Judge (LAJ) evaluation framework achieving 47× cost reduction vs GPT-5 with negligible accuracy loss for automated test coverage assessment
 - Evaluated reasoning LLMs (DeepSeek-R1, Qwen3, GPT-OSS) for sentiment analysis with cross-platform performance benchmarking on NVIDIA H100, NVIDIA GeForce RTX 4090, Apple M3 Max
 - Developed energy efficiency metrics (Energy per Inference, Performance-to-Power Ratio) for sustainable AI deployment
- **Classical ML, Predictive & Risk Modeling Initiatives:**
 - Built financial decisioning pipelines combining classical ML models (Logistic Regression, SVM, Random Forest, XGBoost) with engineered and LLM-derived features — using grid-search hyperparameter tuning, 5-fold cross-validation, sequential feature selection, and log-loss probability calibration to drive risk-sensitive, Kelly-criterion bet sizing
 - Developed ensemble predictive models for ICO / crypto-project success (Naïve Bayes, SVM, Logistic Regression, Random Forest, CatBoost, neural networks) with sentiment- and engagement-based feature engineering
 - Benchmarked the full modeling spectrum — traditional ML (Random Forest, SVM with TF-IDF), deep learning (CNN, RNN, LSTM), and transformers (BERT, RoBERTa) — for fake-news detection and automated maritime risk classification, introducing a novel output-reliability metric, Ratio of Valid Categories
- **Web3 Initiatives:**
 - Developed Mastercard Tokenized Card NFT solution, enabling seamless tokenization of payment cards as Soul Bound Tokens (SBTs) into Web3 wallets
 - Enables users to store credit cards securely as non-transferable NFTs, providing publicly verifiable digital identity for web3 wallet addresses
 - Integrates with crypto exchanges, merchant partners, and payment acquirers through smart contract architecture

- Simplifies Web3 user onboarding by eliminating the need to switch between multiple applications and cumbersome digital wallet funding processes
- Developed DigiPay, a Web3 micropayment stablecoin solution leveraging Layer 2 blockchain and ERC-4337 account abstraction
 - Implemented Digicents smart contract token with high TPS, low-latency blockchain for fast, efficient, and affordable transactions
 - Built smart contract wallet accounts with auto-top-up and auto-withdrawal features, eliminating manual on-ramp/off-ramp processes
 - Integrated with Tokenized Card NFT for secure funding and transaction authentication within the Digicents ecosystem

Aug 2011 – May 2022 Mastercard Asia/Pacific Pte Ltd

Singapore

VP, R&D / Head of Singapore R&D

- Found and lead Singapore R&D team in Mastercard Foundry (previously known as Mastercard Labs) – Mastercard’s global innovation and R&D division; managing a high performing R&D team based in Singapore.
- Work on innovation and R&D in payment/commerce domains across lots of latest and cutting-edge technologies, including but not limited to, EMV Chip technologies, Internet of Things, QR Payments, Connected Car, Humanoid Robots, Conversational AI, AR/VR, Blockchain and Digital Assets, Web3, NFT, DeFi, Metaverse and so on
- Drive complex initiatives from idea through to commercialization hand-over; facilitate the transition of initiatives into Mastercard commercialization process
- As a champion for Technological Innovation, collaborate with various partners (internal departments, external partners, universities) to ensure Mastercard attains a reputation for technological excellence in the Payments space
- **Key Commercialized Projects**
 - Nov 2021: Mastercard partners with city of New Orleans and MoCaFi to announce Crescent City Card program: <https://www.mastercard.com/news/press/2021/november/mastercard-partners-with-city-of-new-orleans-and-mocafi-to-announce-crescent-city-card-program/>
 - Sep 2020: Mastercard partners with Samsung, Airtel and Asante to drive digital inclusion in Africa through Pay-on-Demand services: <https://newsroom.mastercard.com/mea/press-releases/mastercard-partners-with-samsung-airtel-and-asante-to-drive-digital-inclusion-in-africa-through-pay-on-demand-services/>
 - Oct 2018: Mastercard, Centenary Bank and M-KOPA Solar roll-out first of its kind energy solution in Uganda: <https://newsroom.mastercard.com/mea/press-releases/mastercard-centenary-bank-and-m-kopa-solar-roll-out-first-of-its-kind-energy-solution-in-uganda/>
 - Feb 2018: Mastercard Uses Facebook Messenger To Help Small Businesses Go Digital: <https://newsroom.mastercard.com/press-releases/mastercard-uses-facebook-messenger-help-small-businesses-go-digital/>
 - Sep 2017: SPH Buzz Partners With Mastercard To Launch New Hybrid Store Concept: <https://newsroom.mastercard.com/asia-pacific/press-releases/sph-buzz-partners-with-mastercard-to-launch-new-hybrid-store-concept/>
 - Feb 2017: BharatQR – World’s first interoperable QR code acceptance solution launched in India: <https://newsroom.mastercard.com/asia-pacific/press-releases/bharatqr-worlds-first-interoperable-qr-code-acceptance-solution-launched-in-india/>

- May 2016: MasterCard Powers First Commerce Application within SoftBank Robotics' Humanoid Robot "Pepper": <https://newsroom.mastercard.com/press-releases/mastercard-powers-first-commerce-application-within-softbank-robotics-humanoid-robot-pepper/>
- Sep 2015: MasterCard Transforms Aid Distribution (video: <https://youtu.be/W2XIyqeP5w4>): <https://newsroom.mastercard.com/press-releases/mastercard-transforms-aid-distribution/>
- Aug 2012: StarHub unveils 'SmartWallet', in collaboration with DBS, EZ-Link and MasterCard: <https://www.starhub.com/about-us/newsroom/2012/august/starhub-unveils-smartwallet--in-collaboration-with-dbs--ez-link-.html>

Jul 2008 – Aug 2011 PayPal Singapore Development Centre Singapore

Senior Software Engineer / Team Lead

- Lead a team of software developers, working on a pipeline of projects related to Brazil to cater to this fast growing market
- Subject Matter Expert in limits / verification / compliance domain
- **Key Commercialized Projects**
 - Led 12-developer team delivering largest EU compliance project (EU KYB Receiving Limit) in 8 months
 - Built China UnionPay (CUP) integration enabling real-time transaction processing

Oct 2007 – Jul 2008 Streaming21 (Singapore) Pte Ltd Singapore

Software R&D Manager

- Established an R&D centre and led development of Media Relay Server, a leading IPTV streaming server.

Jan 2006 – Sep 2007 Schmidt Electronics (S.E.A.) Pte Ltd Singapore

Leader, Software Development

Feb 2005 – Dec 2005 Unaxis Singapore Pte Ltd Singapore

Software Engineer

Sep 2002 – Feb 2005 ASM Technology Singapore Pte Ltd Singapore

Software Engineer

Aug 1995 – Aug 1998 Beijing Zhida Technology Pte Ltd Beijing, China

Software Engineer

SPEAKING & PANELS

- Invited speaker, IMDA Technical Sharing Session on AI Agents, Singapore (23 April 2026) — alongside Prof Bo An, AAAI Fellow and President's Chair Professor at NTU. <https://www.imda.gov.sg/activities/activities-catalogue/technical-sharing-session-with-ai-agents>
- Invited speaker, InnoAI 2026 Global Summit on Innovating the Future of Artificial Intelligence — listed alongside speakers from JPMorgan Chase, Google, NVIDIA, and eBay. <https://innoaisummit.com/experts>
- Workshop Chair, SENTIRE'25 — 15th Workshop on Sentiment Analysis and Linguistic Linked Data, co-located with IEEE ICDM 2025, Washington DC, November 2025

- Invited speaker, SuperAI Singapore 2025, Marina Bay Sands (19 June 2025) — “Optimizing and Benchmarking Open-Source LLMs for Enterprise Applications.” <https://www.superai.com/superai-sessions/optimizing-and-benchmarking-open-source-llms-for-enterprise-applications>
- Invited panelist, IEEE Conference on Artificial Intelligence (IEEE CAI 2024), Marina Bay Sands, Singapore (June 2024) — panel moderated by Simon See (Director, NVIDIA AI Technology Centre), alongside Prof Masashi Sugiyama (Director, RIKEN Center for Advanced Intelligence Project). 800+ global attendees. https://www.linkedin.com/posts/donghao-huang_invited-to-participate-in-the-ieee-cai-panel-ugcPost-7211296805160968192-e25K/
- Solo session speaker and panelist, CALM SUMMIT '24 (Rasa), October 2024 — “Optimizing and Benchmarking Open-Source LLMs for Conversational AI” plus panel “Benefits of Scalable Language Models.” <https://calmsummit24.com/agenda/>
- Keynote speaker, Token 2049 — WSPN Stable Rise, Singapore (September 2024) — “Building Trust & Seamless Commerce Experience: Leading the Future of Web3 Payments” — presented Mastercard’s Tokenized Card NFT and DigiPay solutions.
- Invited speaker, FCon 2023 Global FinTech Conference, Shanghai, China (2023) — representing Mastercard Labs — “The Future of Commerce in the Web3 Era” (Web3 时代的商业未来). Video: <https://www.infoq.cn/video/3UlpCdtmzxz5OwRjd6j9>
- Invited speaker, Upbit Developer Conference (UDC) 2022, Busan, South Korea (September 2022) — “Mastercard is getting ready for Web3 and the metaverse,” representing Mastercard alongside speakers including Justin Sun (Tron, Founder), Sebastien Borget (The Sandbox, COO & Co-Founder), Bryan Pellegrino (LayerZero Labs, CEO), and Leemon Baird (Hedera/Swirls Labs, Co-CEO). Featured in the official UDC 2022 Report: https://udc-static.dunamu.com/reports/UDC_2022_Report_EN.pdf
- Paper presentations at PAKDD 2024/2025/2026 (Taipei/Sydney/Hong Kong), ICDSAAI 2026 (May 2026, Chengdu, China), AAMAS 2026 (May 2026, Paphos, Cyprus), IEEE ICDM 2025 (Nov 2025, Washington DC), IJCNN 2025 (Jun 2025, Rome, Italy), NAACL 2025 (Apr 2025, Albuquerque, NM), IEEE SSCI 2025 (Mar 2025, Trondheim, Norway), IEEE CAI 2024 (Jun 2024, Marina Bay Sands, Singapore)

AWARDS & RECOGNITIONS

- IMDA Singapore Digital Scholarship (Postgraduate), 2025 — Competitive Singapore government scholarship awarded to senior technology professionals pursuing advanced research
- 1st Prize, SMU Libraries "Generate Your L(AI)brary" Hackathon, 2023 — Won first place at Singapore Management University's AI hackathon (26–27 August 2023). GitHub: https://www.linkedin.com/posts/donghao-huang_yesterday-our-team-won-1st-prize-in-smu-library-ugcPost-7101786441805926400-LW0S/
- Smart Sentosa Challenge Winner, 2016 — Mastercard Asia-Pacific’s Smart City Engine won the Smart Sentosa challenge, a Singapore government smart-city innovation competition

SELECTED MEDIA COVERAGE

- “Engineering the Future of AI: SMU EngD Powers Innovation for a Global Technology Leader” — SMU School of Computing and Information Systems, February 2026. <https://computing.smu.edu.sg/engd/community-stories/engineering-future-ai-smu-engd-powers-innovation-global-technology-leader>

- Token 2049 Side Event Coverage — Investing.com, September 2024. <https://hk.investing.com/analysis/article-110563>
- “支付科技与安全交易” Panel — Sina Finance, March 2018. <https://finance.sina.cn/blockchain/2018-04-02/detail-ifysvnxn1920535.d.html>
- “零距离感受万事达卡的支付黑科技” — Sina / Payment Industry Network, November 2017. https://k.sina.cn/article_6007803312_16617cdb0001001wwd.html
- “不换手机和掏不起钱的人，怎样才能用上新支付技术？” — China E-Banking Network / cnBeta, November 2017. <https://www.cebnet.com.cn/20171127/102444564.html>
- “亞洲最大！萬事達卡支付科技創新，9 間實驗室環腰” — Business Next (數位時代), November 2017. <https://www.bnext.com.tw/article/47137/mastercard-innovation-lab>

VOLUNTEERING EXPERIENCE

Nov 2025 – to date	Bao Yan (Singapore) Education Society	
President		
Jun 2016 – Jun 2020	Peking University Alumni Association (Singapore)	
Vice President		
Jan 2015 – Jan 2019	Startupbootcamp FinTech	Singapore
Mentor		
Mar 2011 – Dec 2015	ConneXions International	Singapore
IT and Payment Consultant		
Nov 2006 – Jul 2008	St. Andrew's Autism Centre	Singapore
Technology Consultant		

TECHNICAL SKILLS

- **AI/ML/GenAI:** LLMs (GPT-4, GPT-5, Claude, Gemini, Llama, Qwen, DeepSeek-R1, Mistral, Magistral, Phi, Nemotron), PyTorch, Transformers, LLM/ML Model Fine-tuning and Optimization, PEFT, LoRA/QLoRA Fine-tuning, LlamaFactory, Unsloth, RAG, LangChain, LangGraph, Autogen, Ollama, LM Studio, Vector Databases, FAISS, Chroma, Prompt Engineering, Few-shot Learning, NLP, Sentiment Analysis, Agentic AI, OpenClaw, NemoClaw, Edge AI Deployment, Energy-efficient Inference
- **GPU/Hardware:** NVIDIA H100, L40, A40, RTX 4080, RTX 4090, RTX A6000, Jetson AGX Orin, DGX Spark; Apple M-series
- **Languages & Frameworks:** Python, C/C++, Java, NodeJS, React, Vue, Spring Boot/Cloud, Solidity
- **Cloud & Infrastructure:** AWS, Azure, Heroku; Microservices, CI/CD, Docker
- **Web3/Blockchain:** Blockchain, Smart Contract, IPFS, NFT, DeFi

PATENTS

(10 granted, 50+ pending — all assigned to Mastercard)

Granted Patents

- **US12417452B2** – Smart chip payment acceptance (2025)
- **US12217246B2** – Method and system for use of an EMV card in a multi-signature wallet for cryptocurrency transactions (2025)
- **US11715100B2** – Electronic system and computerized method for verification of transacting parties to process transactions (2023)
- **US11615406B2** – Method and system for providing a service at a self-service machine (2023)
- **US10535034B2** – Item delivery management systems and methods (2020)
- **US10482499B2** – Method for conducting a transaction (2019)
- **US10423949B2** – Vending machine transactions (2019)
- **US10083427B2** – Method for receiving an electronic receipt of an electronic payment transaction into a mobile device (2018)
- **US9836729B2** – Method and system for conducting a payment transaction and corresponding devices (2017)
- **US8712914B2** – Method and system for facilitating micropayments in a financial transaction system (2014)

Selected Pending Applications

- **WO2026063867A1** – Method and system for seamless cross-network physical-digital (phy-gital) experience (2024)
- **US20260080407A1** – Method and system for seamless cross-network physical-digital (phy-gital) experience (2024)
- **WO2024058723A1** – Digitization of payment cards for web 3.0 and metaverse transactions (2022)
- **US20230059546A1** – Access control system (2021)
- **EP4356332A2** – Method and system for mediated cross ledger stable coin atomic swaps using hashlocks (2021)
- **US20210406887A1** – Method and system for merchant acceptance of cryptocurrency via payment rails (2020)
- **US20210158316A1** – Electronic system and computerized method for controlling operation of service devices (2019)
- **US20200175489A1** – Method and system for QR code originated vending (2017)
- **WO2017200643A1** – Method and system for allocating and transacting with multiple non-financial commodities (2016)
- **US20180025332A1** – Transaction facilitation (2016)

... and 40+ additional pending patent applications across payment technology, blockchain, AI, and IoT domains.

RESEARCH PAPERS

(papers at top-tier venues are highlighted in yellow; full texts are accessible via the hyperlinked paper titles)

Artificial Intelligence

1. Huang, D., Drietomský, T., Barrett, B. & Wang, Z. (2026). [How Small Can You Go? LoRA Fine-Tuning 270M–8B Models for Merchant Information Extraction in Financial Transactions](#). Accepted for publication in 2026 IEEE International Conference on Data Mining (ICDM 2026), Shenyang, China.
2. Teo, N., Huang, D. & Wang, Z. (2026). [A Survey of Commonsense Reasoning in LLMs](#). ACM Computing Surveys, vol. 58, no. 14, pp. 1–34 (2026).
3. Chua, J. K., Huang, D. & Wang, Z. (2026). [A Novel Hierarchical Multi-agent System for Payments Using LLMs](#). In *Advances in Knowledge Discovery and Data Mining. PAKDD 2026. Lecture Notes in Computer Science*, vol. 16598, pp. 28–39. Springer, Singapore.
4. Huang, D., & Wang, Z. (2026). [Task Complexity Matters: An Empirical Study of Reasoning in LLMs for Sentiment Analysis](#). In *Advances in Knowledge Discovery and Data Mining. PAKDD 2026. Lecture Notes in Computer Science*, vol. 16600, pp. 288–299. Springer, Singapore.
5. Huang, D., Chua, J. K., & Wang, Z. (2027). [Beyond Task Success: Measuring Workflow Fidelity in LLM-Based Agentic Payment Systems](#). In *Trends and Applications in Knowledge Discovery and Data Mining. PAKDD 2026. Lecture Notes in Computer Science*, vol. 16603, pp. 357–362. Springer, Singapore.
6. Huang, D., Pai, P., & Wang, Z. (2027). [Entity Matching with LLMs at Scale: Accuracy, Cost, and Reasoning Effects](#). In *Trends and Applications in Knowledge Discovery and Data Mining. PAKDD 2026. Lecture Notes in Computer Science*, vol. 16603, pp. 112–124. Springer, Singapore.
7. Huang, D., Chew, S., & Wang, Z. (2027). [CMAS4G2: A Locally Deployed Cooperative Multi-agent System for Gherkin Acceptance Test Generation](#). In *Trends and Applications in Knowledge Discovery and Data Mining. PAKDD 2026. Lecture Notes in Computer Science*, vol. 16603, pp. 156–167. Springer, Singapore.
8. Huang, D. & Pandey, M. (2026). [An Agentic Voice-Based Assistant for Interactive Conversation and Guidance in Real-World Environments](#). In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '26)*, Paphos, Cyprus, pp. 4167–4169. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
9. Chua, J. K., Tse, C. H., Huang, D., & Wang, Z. (2026). [ConvPayMAS: Conversational Payment Multi-Agent System with Agent-to-Agent Protocol and Three-Mandate Verification](#). In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '26)*, Paphos, Cyprus, pp. 4071–4073. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
10. Huang, D., Malwe, G., & Wang, Z. (2026). [When Agents Fail to Act: A Diagnostic Framework for Tool Invocation Reliability in Multi-Agent LLM Systems](#). 2026 9th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 492–497. IEEE.
11. Huang, D., Chew S., Dutkiewicz A., & Wang, Z. (2026). [LLM-as-a-Judge for Scalable Test Coverage Evaluation: Accuracy, Operational Reliability, and Cost](#). Presented at *AAAI 2026 Workshop on Next-Gen Code Development with Collaborative AI Agents*.
12. D. Huang, N. Teo, A. Chekuru, E. Deutschman and Z. Wang, "[Deploying Reasoning LLMs for Sentiment Analysis: Architecture Trade-offs on Consumer Hardware](#)," 2025 IEEE International Conference on Data Mining Workshops (ICDMW), Washington, DC, USA, 2025, pp. 2048-2055.
13. S. Samuel, D. Huang, J. T. Y. Hong and Z. Wang, "[Integrating LLM Sentiment Analysis into Machine Learning for Soccer Betting](#)," 2025 IEEE International Conference on Data Mining Workshops (ICDMW), Washington, DC, USA, 2025, pp. 2156-2164.
14. D. Huang and Z. Wang, "[Explainable Sentiment Analysis With DeepSeek-R1: Performance, Efficiency, and Few-Shot Learning](#)," in *IEEE Intelligent Systems*, vol. 40, no. 6, pp. 52-63, Nov.-Dec. 2025.
15. Huang, D., Ng, D., Pen, H., Wang, Z., & Cambria E. (2025). [Evaluating the Impact of LLM-Manipulated Content on Fake News Detection](#). In: Yuan, S., Malliaros, F., Zheng, X. (eds) *Trends and Applications in Knowledge Discovery and Data Mining. PAKDD 2025. Lecture Notes in Computer Science*, vol. 15835, pp. 375–386. Springer, Singapore.

16. Teo, N., Huang, D., Cambria, E., Wang, Z. (2025). [Large Language Models for Logical Fallacy Detection](#). In: Yuan, S., Malliaros, F., Zheng, X. (eds) *Trends and Applications in Knowledge Discovery and Data Mining, PAKDD 2025*. Lecture Notes in Computer Science, vol. 15835, pp. 387–398. Springer, Singapore.
17. D. Huang and Z. Wang, "[LLMs at the Edge: Performance and Efficiency Evaluation with Ollama on Diverse Hardware](#)," *2025 International Joint Conference on Neural Networks (IJCNN)*, Rome, Italy, 2025, pp. 1-8.
18. Huang, D., Nguyen, T. S., Liausvia, F., & Wang, Z. (2025, April). [RAP: A Metric for Balancing Repetition and Performance in Open-Source Large Language Models](#). In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)* (pp. 1479-1496).
19. Huang, D., Malwe, G., Varshney, R., Tse, A., & Wang, Z. (2025, March). [Scalable Invoice Reconciliation for SMEs: Edge-Deployed LLMs in Multi-Agent Systems](#). In *2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India*(pp. 1-6). IEEE.
20. D. Huang and Z. Wang, "[Optimizing Chinese-to-English Translation Using Large Language Models](#)," *2025 IEEE Symposium on Computational Intelligence in Natural Language Processing and Social Media (CI-NLPSoMe)*, Trondheim, Norway, 2025, pp. 1-7.
21. D. Huang and Z. Wang, "[Logical Reasoning with LLMs via Few-Shot Prompting and Fine-Tuning: A Case Study on Turtle Soup Puzzles](#)," *2025 IEEE Symposium on Computational Intelligence in Natural Language Processing and Social Media (CI-NLPSoMe Companion)*, Trondheim, Norway, 2025, pp. 1-5.
22. Wang, Z., Huang, D., Cui, J. et al. (2025). [A review of Chinese sentiment analysis: subjects, methods, and trends](#). *Artif Intell Rev* 58, 75 (2025).
23. Huang, D., Fu, X., Yin, X., Pen, H., & Wang, Z. (2024). [Automating maritime risk data collection and identification leveraging large language models](#). In *2024 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 433–439). IEEE.
24. Huang, D., Hu, Z., & Wang, Z. (2024, June). [Performance Analysis of Llama 2 Among Other LLMs](#). In *2024 IEEE Conference on Artificial Intelligence (CAI)* (pp. 1081-1085). IEEE.
25. Huang, D., & Wang, Z. (2024, April). [Evaluation of Orca 2 Against Other LLMs for Retrieval Augmented Generation](#). In: Wang, Z., Tan, C.W. (eds) *Trends and Applications in Knowledge Discovery and Data Mining, PAKDD 2024*. Lecture Notes in Computer Science, vol. 14658, pp. 3–19. Springer, Singapore.
26. Huang, D., Samuel, S., Hyunh, Q. T., & Wang, Z. (2024, April). [From Tweets to Token Sales: Assessing ICO Success Through Social Media Sentiments](#). In: Wang, Z., Tan, C.W. (eds) *Trends and Applications in Knowledge Discovery and Data Mining, PAKDD 2024*. Lecture Notes in Computer Science, vol. 14658, pp. 57–69. Springer, Singapore.

Theoretical Physics

27. Lin, W. B., Huang, D. H., Zhang, X., & Brandenberger, R. (2001). [Nonthermal production of weakly interacting massive particles and the subgalactic structure of the universe](#). *Physical Review Letters*, 86(6), 954.
28. Huang, D. H., Lin, W. B., & Zhang, X. M. (2000). [Remark on approximation in the calculation of the primordial spectrum generated during inflation](#). *Physical Review D*, 62(8), 087302.