T5 -pretrained language models (PLMs) have made significant progress on various NLP tasks recently . but they face challenges when it comes to domain-specific tasks such as legal AI . these tasks often involve intricate expertise, expensive data annotation, and limited training data availability . we propose a human-oriented artificial–natural parallel system for organized intelligence (HANOI)-Legal based on the parallel learning framework .

Bart-

Pretrained language models (PLMs) have made significant progress on various NLP tasks recently. PLMs encounter challenges when it comes to domain-specific tasks such as legal AI. These tasks often involve intricate expertise, expensive data annotation, and limited training data availability. We propose a human-oriented artificial–natural parallel system for organized intelligence (HANOI)-Legal based on the parallel learning (PL) framework. Our experiments in two legal domain tasks show that HANOi-Legal achieved an excellent performance in low-resource scenarios compared to the state-of-the-art prompt-based approach. It leverages PLMs’ linguistic capabilities acquired from a variety of open datasets to generate task-specific models. By combining a scalable unified-task-format reformulation and a unified-prompt-based training pipeline, HANNOI-Legal leverages the PLM’s linguistic capabilities from diverse open datasets for these tasks. It also labels a few task- specific legal data through distributed autonomous operations (DAO-II) for further fine-tuning.

Gemini -

HANOI-Legal, a human-oriented artificial-natural parallel system for legal AI, tackles domain-specific NLP challenges with limited data. Using a parallel learning framework, it combines a large-scale PLM with a unified prompt-based training approach (UniPrompt) and task-specific data sourced through distributed autonomous operations (DAO-II). This allows HANOI-Legal to leverage the linguistic capabilities of PLMs for legal tasks with limited resources, achieving state-of-the-art results.

GPT-4 -In recent years, the growing field of Natural Language Processing (NLP) has witnessed significant strides, particularly with the advent of pretrained language models (PLMs). However, these models often struggle with domain-specific tasks, such as those in legal AI, due to the complexity of the domain, costly data annotation, and scarcity of training data. To address these challenges, a novel approach termed the Human-Oriented Artificial–Natural Parallel System for Organized Intelligence (HANOI)-Legal has been introduced, utilizing the parallel learning (PL) framework.

Previous research highlights significant advances in PLMs, but the adaptation to specialized fields like legal AI remains arduous. The HANOI-Legal model innovatively incorporates a two-fold strategy. Initially, it sets up an artificial system using a PLM pre-trained on a large-scale corpus, establishing a robust foundation. Subsequently, to tailor the PLM for legal-specific tasks with limited resources, the model introduces UniPrompt—a unified prompt-based training framework. This framework is designed to harness diverse open datasets effectively, enhancing the model's adaptability and performance in legal contexts.

Numerous studies have addressed this topic with varying approaches, yet the HANOI-Legal model stands out by integrating task-specific data labeled through Distributed Autonomous Operations (DAO-II) for further refinement. This integration facilitates the fine-tuning process, crucial for the model's accuracy in domain-specific applications. By merging a scalable unified-task-format reformulation with a unified-prompt-based training pipeline, HANOI-Legal capitalizes on the linguistic capabilities of PLMs derived from various datasets, crafting models that are finely tuned for specific legal tasks.

The efficacy of HANOI-Legal has been empirically validated through experiments on two legal domain tasks, where it demonstrated superior performance in low-resource settings compared to existing state-of-the-art prompt-based methods. This breakthrough underscores the potential of parallel learning frameworks in overcoming the limitations posed by specialized domains, paving the way for more sophisticated and resource-efficient AI tools in the legal field.

GPT-3

The research introduces a novel approach, HANOI-Legal, to address the challenges faced by pretrained language models (PLMs) in domain-specific tasks like legal AI. The proposed system leverages a parallel learning (PL) framework, utilizing a large-scale corpus for pretraining and a unified prompt-based training framework called UniPrompt to adapt PLMs to legal tasks with limited resources. Through distributed autonomous operations (DAO-II), task-specific legal data is labeled for fine-tuning. By combining a unified-task-format reformulation and a unified-prompt-based training pipeline, HANOI-Legal effectively harnesses PLMs' linguistic capabilities from diverse open datasets to generate task-specific models. Experimental results in two legal domain tasks demonstrate that HANOI-Legal outperforms state-of-the-art prompt-based approaches in low-resource scenarios. This innovative approach not only overcomes the challenges of intricate expertise, expensive data annotation, and limited training data availability in legal AI but also showcases the potential of parallel learning for enhancing performance in domain-specific tasks. The study highlights the importance of a human-oriented artificial–natural parallel system for organized intelligence in advancing the capabilities of PLMs for specialized applications like legal intelligence.