Fine-grained Contract NER using instruction based model

Hiranmai Sri Adibhatla*, **Pavan Baswani***, **Manish Shrivastava**
Language Technologies Research Center, KCIS, IIIT Hyderabad, India.
{hiranmai.sri, pavan.baswani}@research.iiit.ac.in
m.shrivastava@iiit.ac.in

Abstract

Lately, instruction-based techniques have made significant strides in improving performance in few-shot learning scenarios. They achieve this by bridging the gap between pre-trained language models and fine-tuning for specific downstream tasks. Despite these advancements, the performance of Large Language Models (LLMs) in information extraction tasks like Named Entity Recognition (NER), using prompts or instructions, still falls short of supervised baselines. The reason for this performance gap can be attributed to the fundamental disparity between NER and LLMs. NER is inherently a sequence labeling task, where the model must assign entity-type labels to individual tokens within a sentence. In contrast, LLMs are designed as a text generation task. This distinction between semantic labeling and text generation leads to subpar performance.

In this paper, we transform the NER task into a text-generation task that can be readily adapted by LLMs. This involves enhancing source sentences with task-specific instructions and answer choices, allowing for the identification of entities and their types within natural language. We harness the strength of LLMs by integrating supervised learning within them. The goal of this combined strategy is to boost the performance of LLMs in extraction tasks like NER while simultaneously addressing hallucination issues often observed in LLM-generated content. A novel corpus Contract NER comprising seven frequently observed contract categories, encompassing named entities associated with 18 distinct legal entity types is released along with our baseline models. Our models and dataset are available to the community for future research[1].

1 Introduction

Contracts are legally enforceable agreements that outline the rights and responsibilities of involved parties, governing interactions between companies, employees, contractors, customers, and suppliers. In contrast to the corpora commonly utilized for pre-training deep models, the composition and terminology of contracts differ significantly. Contracts usually adhere to specific template formats to ensure unambiguity. Given the significance of precise word selection and sentence structure in legal documents, even minor ambiguities can result in unintended interpretations and consequences. Therefore, meticulous drafting and thorough review of contracts are crucial, as they serve as essential instruments for managing business relationships and mitigating risks. Creating automated tools and applications can play a crucial role in diminishing the time required to accomplish contract understanding, drafting, and review.

Among the tasks proposed to facilitate contract review, entity extraction which is based on named entity recognition (NER) plays a fundamental role in extracting information and processing the contract. Commonly, systems designed to recognize named entities identify individuals, organizations, dates, locations, currency terms, and more. However, named entities found in legal texts exhibit nuanced differences and demand a more fine-grained analysis. Extracting named entities or contract elements manually can be time-consuming, expensive, and repetitive, prompting the demand for automation sought by legal professionals and their clients. With this in mind, our paper aims to address the challenge of automatically identifying crucial contract elements. These elements, which include parties involved, specific dates, monetary values, explicit rights and obligations, and relevant governing laws, hold significant importance within a contract. By automating the identification of these elements,

* Authors contributed equally

[1] https://github.com/pavanbaswani/ContractNER/

we can streamline the contract analysis process, reduce costs, and improve overall efficiency in the legal domain. The paper uses the terms "Named Entity Recognition" and "Contract Element Extraction" interchangeably in the context of contracts.
Previous studies have focused on identifying fine-grained named entities in judgment documents (Kalamkar et al., 2022; Leitner et al.; Barriere and Fouret, 2019). However, when it comes to contracts, similar efforts have been constrained by the limited coverage of entity types (Au et al., 2022; de Almeida et al., 2020) and contract categories (Leivaditi et al., 2020; Niklaus et al., 2023). This paper presents the development of a prompt-based corpus for contract Named Entity Recognition (NER) encompassing eighteen fine-grained entity types from seven commonly encountered contract types. The study includes the creation of baseline models for sequence labeling, parameter-efficient learning, and prompt-based learning using LLMs and involves a comparative analysis of LLMs performance in information extraction tasks. Additionally, the guidelines used for the construction of the corpus are presented in detail in this work.

2 Related Work

Over the past few years, there has been a notable rise in research activity centered around document and text processing in the legal domain. This surge in interest has led to the development of numerous datasets, tasks, and applications, including but not limited to prior case retrieval (Al-Kofahi et al., 2001; Jackson et al., 2003), summarization (Bhattacharya et al., 2019), events and named entities extraction (Kalamkar et al., 2022; Lagos et al., 2010), and judgment prediction (Xiao et al., 2018; Chalkidis et al., 2019; Malik et al., 2021).
A considerable amount of work has been done in contract analysis and information extraction from contracts (Yang et al., 2013; Silva et al., 2020; Mittal et al., 2015). Extracting contract elements from legal documents, such as contracts, has been a long-standing challenge for the legal and artificial intelligence (AI) communities.
One of the earliest approaches to contract element extraction involved using rule-based methods, where experts manually created rules with hand-crafted features, word embeddings, and part-of-speech tag embeddings to identify and extract specific elements from contracts (Chalkidis et al., 2017). The primary emphasis is placed on the extraction of essential elements, including the contract's title, start, and termination dates, contracting parties, contract value, and more. Although these methods demonstrated effectiveness, they were limited by their inability to handle the complexity and variability of natural language. This is attributed to the use of linear classifiers, which may not be able to capture complex relationships between the input features and output labels.
Use of deep learning methods for contract element extraction (Chalkidis and Androutsopoulos, 2017) along with conditional Random Fields (Finkel et al., 2005; Xu and Sarikaya, 2013) was popular for sequence labeling tasks prior to neural networks. The authors proposed a Bidirectional LSTM (BILSTM) model that operates on word, part-of-speech (POS) tags, and token shape embeddings. This model was tested against the linear sliding-window classifiers (Chalkidis et al., 2017). The advantages of this approach are that it does not rely on manually written rules and it can handle multi-token contract elements.
Recently, neural networks (Huang and Xu, 2015; Ma and Hovy, 2016; Chalkidis et al., 2019) based approaches were employed for contract elements extraction. This approach (Chalkidis et al., 2019) investigates the task of contract element extraction and compares the performance of several neural network-based models such as LSTM-based encoders, and transformers. BERT (Devlin et al., 2019) based approaches (Zhang et al., 2020; Chen et al., 2019) formulate contract elements as a sequence labeling task by adapting BERT (Zhang et al., 2020), a state-of-the-art language model. These models show promising results in extracting important content elements from business documents. The study demonstrates that even with a modest amount of annotated data, the model can achieve reasonable accuracy, which is valuable for practical applications. Prompt-based methods (Barriere and Fouret, 2019) have shown significant progress in few-shot learning scenarios by bridging the gap between language model pre-training and fine-tuning for downstream tasks.
The majority of the aforementioned approaches are constrained by their focus on only one or two types of business documents, which hinders their ability to generalize well to other types of business documents. The specificity of the training data and features used in these approaches may

not fully encompass the diverse characteristics and structures present in different business document domains. Therefore, there is a need to develop more versatile methods that can effectively handle a wider range of business document types, ensuring broader applicability and adaptability in contract analysis tasks. Our work is inspired by these approaches, we build and release an annotated dataset along with train language models to enable the automatic identification and extraction of contract elements from contracts. The improvements presented in this paper include significant coverage in the targeted document types and fine-grained classification for contract elements. This work aims to facilitate the creation of various legal AI applications that can automatically detect fine-grained named entities from contracts. This advancement aims to streamline contract processing and reviewing, making the entire process much more efficient and user-friendly.

3 Contract NER Corpus

Our dataset comprises diverse legal contracts sourced from SEC EDGAR[2](Electronic Data Gathering, Analysis, and Retrieval). It is a comprehensive online database maintained by the U.S. Securities and Exchange Commission (SEC). It serves as a centralized repository for a wide range of financial and business-related documents submitted by publicly traded companies, investment firms, and other entities regulated by the SEC.

3.1 Contract Preprocessing

In the process of curating our legal contract dataset and making the contracts amenable to further analysis, we extracted plain text from raw documents sourced. In the scraped dataset, we observed a diverse range of contract titles, but not all titles were equally represented. To address this imbalance, we employed heuristics to extract the most common contract titles based on their frequency of occurrence. Table 1 outlines the contract titles extracted and the counts of contracts extracted for each title. By including a diverse set of frequently titled documents in the training data, the model gains a deeper understanding of legal contract structures and their terminology. Leveraging a rich and varied training dataset enables our model to become a powerful tool for handling contract-related entity extraction, and streamlining contract analysis efficiently with precision and efficiency.

Contract Type	Train	Dev	Test
Employment	113	19	15
Credit	4	2	2
Purchase	53	6	6
Loan	15	4	4
Lease	35	4	4
Indemnification	21	2	2
Consulting	16	2	2

Table 1: Legal Contract types and their documents' distribution

3.2 Pre Annotations

Manual annotation using entity recognition taggers is a crucial and labor-intensive process. It involves human annotators carefully examining the text data and marking specific words or phrases that represent named entities, such as names of people, organizations, locations, and other proper nouns. Entity taggers are NLP tools that extract mentions of entities (such as people, places, or objects of interest) from a document. They are used for various purposes including information extraction, and question-answering. Different entity recognition taggers are available based on their purpose and scope. General-purpose taggers are versatile annotation tools used for various tasks, such as classification, span detection, entity tagging, and part-of-speech tagging. Some commonly used tools for generic tagging tasks include GATE Teamware (Bontcheva et al., 2013), NameTag (Straková et al., 2014), SELECTIVE ANNOTATION (Do Dinh et al., 2015), SLATE (Kummerfeld, 2019), and DoTAT tool (Lin et al., 2022). They were largely utilized to perform the generic tagging tasks mentioned above. On the other hand, there are named entity taggers like WebAnno (Yimam et al., 2013), (Yimam et al., 2014), Open Annotation (OA)(Pyysalo et al., 2015), TALEN(Mayhew and Roth, 2018), APLenty (Nghiem and Ananiadou, 2018), AlpacaTag (Lin et al., 2019), CroAno (Zhang et al., 2021), Doccano (Nakayama et al., 2018), Label Studio (Tkachenko et al., 2020-2022) and INCEpTION (Klie et al., 2018) that work well with entity tagging. However, some of these taggers are not open-sourced, and few lack support for preloaded annotations using available entity taggers

[2] https://www.sec.gov/edgar/search-and-access

like Spacy[3] or LexNLP[4]. To enhance the annotation process and enable pre-annotations from available pre-trained models, we used in-house named entity tagger that can serve our specific purposes effectively. We leverage the few-shot predictions capability of ChatGPT[5] and predictions from LexNLP to auto-populate annotations related to predefined entity categories. Entity extraction using ChatGPT involves providing context from the contract and posing an instruction. An example of the prompt we used is in Figure 1. The model's few-shot capability enables it to extract various entities, such as dates, parties, acts, governing laws, and amounts, from the contract. This flexibility and adaptability make it a valuable tool for automated analysis of legal documents. For other specific entities like generic dates, addresses, courts, and acts, we utilize the LexNLP python library, which employs trained models, heuristics, and dedicated functions to identify and extract entities. For example, LexNLP offers a function to extract generic dates by scanning the input text and retrieving all date-related entities. The output of the extraction process presents well-structured representations of the identified entities, typically in lists or dictionaries, ready for further processing or analysis to meet the application's specific requirements.

The choice to utilize both ChatGPT and LexNLP for pre-annotations serves the purpose of optimizing accuracy and ensuring precise matches, as well as broadening the scope of annotations. For a few entities including *Act* and *Regulation*, LexNLP tends to have a higher exact match rate compared to ChatGPT. However, when it comes to distinguishing and labeling dates as *Effective*, *Termination*, or *Renewal*, ChatGPT proves to be more adept. This combined approach empowers efficient and accurate entity extraction, streamlining the annotation and analysis of legal contracts.

3.3 Manual Review and Corrections

In this paper, we also emphasize the significance of guidelines in human-annotated tasks. Guidelines play a vital role in reducing ambiguity and ensuring creation of accurate datasets. Specifically, we focused on contract-specific entities commonly found within contracts and formulated detailed guidelines to facilitate the annotation process.

[3] https://spacy.io/

[4] https://github.com/LexPredict/lexpredict-lexnlp

[5] https://chat.openai.com/

To ensure consistency and reliability in the annotation, we enlisted three groups, each comprising three student annotators. The annotators were provided with the guidelines and tasked with annotating the contract-specific entities in the given pre-annotated contracts. The annotations produced by the annotators were thoroughly evaluated against the guidelines. Based on the understanding of the task and the quality of their annotations, two annotators were selected for further analysis. Each of the chosen annotators was assigned 125 contracts to annotate, allowing us to assess the consistency and precision of their annotations on a substantial sample size. The annotators were provided with comprehensive guidelines for each entity category. We list the guidelines in Appendix A Table 5. Additionally, they were given general guidelines to ensure consistency and accuracy in their annotations:

1. **Concise:** The span marking the entity should be succinct and directly relevant to the entity's representation in the contract.

2. **Correct:** While there might be multiple valid options for the entity type in the contract, annotators were instructed to select the most appropriate and accurate answer based on the context and guidelines provided.

The annotators were given access to the pre-annotated sentences (section 3.2). Their task involved either rectifying incorrect annotations or adding any missing annotations as necessary. Figure 2 and Figure 3 show examples of annotators modifying and correcting the pre annotations.

The objective of enhancing the accuracy and reliability of our dataset was accomplished by incorporating these comprehensive guidelines and consistently conducting sampling and evaluation of the annotated data.

4 Experiments and Analysis

In this section, we detail the exhaustive experiments on fine-grained named entities found in contracts and verify the effectiveness of instruction models for named entity recognition tasks. Although LLMs (Large Language Model-based Models) have achieved remarkable success in various NLP tasks like text generation, summarization, and sentiment analysis, their performance in information extraction tasks, particularly in Named Entity Recognition (NER), is still lacking compared to

```
PROMPT = (
    "Entity Definition:\n"
    "1. ContractTitle: Short name or full name of the contract document.\n"
    "2. ContractParties:  Names of the two or more parties who signed the contract.\n"
    "3. EffectiveDate: The date from when the contract is effective.\n"
    "4. SalaryCompensation: Salary or compensation mentioned for the employee in Employee type Agreements.\n"
    "5. GoverningLaw: The state/country's law that governs the interpretation of the contract.\n"
    "6. EmploymentRole: The role for which an employee is employed.\n"
    "\n"
    "Output Format:\n"
    """{{"ContractTitle": [list of entities present], "ContractParties": [list of entities present], "EffectiveDate": [list of entities present],"SalaryCompensation": [list of entities present],
    "GoverningLaw": [list of entities present],"EmploymentRole": [list of entities present]}}\n"""
    "If no entities are presented in any categories keep it None\n"
    "\n"
    "Examples:\n"
    "\n"
    """1. Sentence: THIS PURCHASE AND SALE AGREEMENT (this "Agreement") made as of the ____ day of October, 2015 between AMERCO REAL ESTATE COMPANY, a Nevada corporation, having an address at 2727
    North Central Avenue, Phoenix, Arizona 85004 ("Seller") and 23RD AND 11TH ASSOCIATES, L.L.C., a Delaware limited liability company, having an address c/o The Related Companies, L.P., 60 Columbus
     Circle, New York, New York 10023.\n"""
    """Output: {{"ContractTitle": ["PURCHASE AND SALE AGREEMENT"], "ContractParties": ["AMERCO REAL ESTATE COMPANY","23RD AND 11TH ASSOCIATES, L.L.C"], "EffectiveDate": ["None"],
    "SalaryCompensation":["None"],"GoverningLaw":["None"],"EmploymentRole":["None"]}}\n"""
    "\n"
    """2. Sentence: TO EXECUTIVE EMPLOYMENT AGREEMENT ("Amendment") is entered into as of August 8, 2016, to be effective as of July 1, 2016, by and between Aqua Metals, Inc., a Delaware corporation
    ("Company"), and Thomas Murphy ("Executive").\n"""
    """Output: {{"ContractTitle": ["EXECUTIVE EMPLOYMENT AGREEMENT"], "ContractParties": ["Aqua Metals, Inc.","Thomas Murphy"], "EffectiveDate": ["July 1, 2016"],"SalaryCompensation":["None"],
    "GoverningLaw":["None"],"EmploymentRole":["None"]}}\n"""
    "\n"
    """3. Sentence: The Employee will be paid an annual salary of Three Hundred Eighty Thousand Dollars ($380,000).\n"""
    """Output: {{"ContractTitle": ["None"], "ContractParties": ["None"], "EffectiveDate": ["None"],"SalaryCompensation":["$380,000"],"GoverningLaw":["None"],"EmploymentRole":["None"]}}\n"""
    "\n"
    """4. Sentence: 2.7.Applicable Law.  This Guaranty shall be construed in accordance with and governed by the laws of the State of New York, without regard to conflict of laws principles.\n"""
    """Output: {{"ContractTitle": ["None"], "ContractParties": ["None"], "EffectiveDate": ["None"],"SalaryCompensation":["None"],"GoverningLaw":["State of New York"],"EmploymentRole":["None"]}}\n"""
    "\n"
    """5. Sentence: Employee serves as its Executive Vice President & Chief Commercial Banking Officer responsible for managing and coordinating the Bank's commercial banking activities.\n"""
    """Output: {{"ContractTitle": ["None"], "ContractParties": ["None"], "EffectiveDate": ["None"],"SalaryCompensation":["None"],"GoverningLaw":["None"],"EmploymentRole":["Executive Vice President
    & Chief Commercial Banking Officer"]}}\n"""
    "\n"

    """6. Sentence: {}\n"""
    """Output: """
)
```

Figure 1: Prompt For Few-Shot Learning in ChatGPT

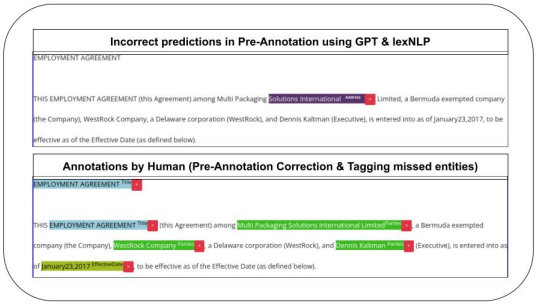

Figure 2: Add Missed Annotation

Figure 3: Rectify Pre Annotation

supervised approaches. Additionally, LLMs encounter the issue of hallucination, which limits their usability in critical information retrieval tasks, where accuracy is crucial. To overcome these limitations, a promising approach is to harness the strengths of both LLMs and supervised models through a combination strategy. When fine-tuned on NER-specific data, LLMs can effectively learn to recognize and extract named entities, surpassing the zero-shot and few-shot capabilities of LLMs.

4.1 Models

In our experiments, we compare popular NER model architectures including prompt-based methods.

1) Sequence labeling models: We apply the traditional sequence labeling method for named entity recognition with the token classification method of BERT (Devlin et al., 2019). We extend BERT (LEGAL-BERT-BASE) for sequence labeling in order to identify phrases of interest. It enables fine-grained entity recognition at the token level, allowing for precise localization and classification of entities.

2) Parameter Efficient models: Parameter-efficient models (Liu et al., 2022) have become increasingly popular in recent times. These models focus on updating only a small subset of parameters during the adaptation of a pre-trained model to downstream tasks. A notable example of parameter-efficient tuning is Low-Rank Adaptation (LoRA) (Hu et al., 2021), which aims to reduce the number of trainable parameters by employing low-rank representations. We fine-tune our dataset with the token classification method of roberta-large (Liu et al., 2019) model. LoRA was applied on the large model to attain efficiency in storage and training. With significantly fewer parameters, LoRA allows for a more streamlined and resource-efficient model, making it a favorable option.

3) Prompt based models Having observed the benefits of few-shot learning in our pre-annotations, we decided to explore the potential of prompt-based models, which have gained significant importance in the field. These models reframe the sequence labeling task as a generation problem, providing

a fresh perspective to tackle the NER task. To align our dataset with this innovative approach, we transformed it into an instruction-based generative framework inspired by NER model based on instructions (Wang et al., 2022). By combining source sentences with descriptive task instructions and limited answer options, we crafted a setup that enhances the model's ability to understand and generate relevant entities. Finally, we fine-tuned the T5-small model (Raffel et al., 2020) on this modified dataset, capitalizing on the power and versatility of prompt-based learning to further improve our NER results. We opted for T5-small due to its architecture, which includes both encoder and decoder components. Information extraction tasks tend to benefit from architectures that incorporate both encoder and decoder, as opposed to models that only feature a decoder. The combination of prompt-based techniques and T5-small fine-tuning improved the performance of our NER system.

4.2 Experimental Setup

Hyper Parameters: For training the model, we utilized consistent hyperparameters, such as a sequence length of 512, a learning rate of 5e-5, Adam optimizer, and a batch size of 4, while also employing the number of beams to 3 for the Prompt based Model. The training was performed on a machine with the following hardware specifications: Nvidia P100 GPU with 16GB memory, operating at 1.32GHz GPU clock, supported by 2 CPU cores and 12GB RAM, all hosted on the Kaggle platform.

4.3 Data Induction

The data annotation and induction process was conducted in three stages, each with varying token samples, and the model's performance was evaluated on unseen inference data picked from SEC. The contract categories and titles gathered from the SEC were vast, making it impractical to train and annotate all titles at once. Therefore, we conducted a staged induction experiment. This approach aimed to assess the model's ability to generalize to the language used in contracts and its adaptability to new, similar data. The results of this experiment, when applied to unknown data, provide the confidence that the model can effectively handle unseen contract categories.

During the first stage, the model was trained exclusively on contracts pertaining to the *Employment* category. In the second stage, a few related contract categories *Credit*, *Lease*, were added, and the model was retrained. Subsequently, in the third stage, additional diverse contract categories *Consulting*, *Loan*, and *Indemnification* were included, and the model underwent retraining.

All contracts are divided into paragraphs since a paragraph as a unit might be of a higher value than an isolated sentence. Table 2 denotes the number of paragraphs and unique tokens observed in each stage. Table 3 denotes the fine-grained contract entity-wise distribution across each stage.

Our instruction-based T5-small model's performance in all three stages was then evaluated on the unseen inference data. The primary objective is to assess its proficiency in dealing with novel categories. Visual representations in Appendix B Figure 4, Figure 5 and Figure 6 showcase the inferred results using our in-house annotation tool. It is evident that the model trained on Stage1 data struggled to recognize entities absent from its training data. Conversely, models trained on Stage2 and Stage3 exhibit enhanced generalization capabilities and excel in recognizing patterns and entities beyond their training data, reflecting their real-world applicability.

4.4 Results

Table 4 presents the outcomes achieved with three baseline models: sequence labeling token classification, a parameter-efficient model fine-tuned on a large pre-trained language model coupled with LoRA, and an instruction-based model fine-tuned using T5-small coupled with LoRA on our complete dataset. Our observations indicate that instruction-based models have outperformed both sequence labeling and parameter-efficient models. This outcome supports our hypothesis that supervised learning on large language models (LLMs) leads to improved accuracy. In the case of a few entity categories such as *Rent* and *Shares*, where token-based classification in both sequence-based models and parameter-efficient models failed to produce results on the test data due to limited samples for those categories, prompt-based models demonstrate superior performance. This underscores the importance of using thoughtfully crafted prompts to direct the model towards generating accurate responses or accomplishing specific tasks, especially in scenarios where data is scarce. The same principle extends to other entities, where we notice higher precision and recall values. Tasks that encompass a diverse set of inputs and outputs

	Stage-1			Stage-2			Stage-3		
	Train	**Dev**	**Test**	**Train**	**Dev**	**Test**	**Train**	**Dev**	**Test**
# Tokens	243706	29801	23167	384439	68084	78427	531293	92685	92543
# Unique tokens	10813	4091	3710	17777	7405	8936	25464	9592	10134
# Paras	2986	327	267	4744	770	929	6882	1059	1113
Avg para length	81.61	90.81	86.51	81.02	88.07	84.25	77.2	87.45	83.14
Max para length	641	656	542	947	1557	2452	2725	1557	2452

Table 2: Data distribution Statistics

	Stage-1			Stage-2			Stage-3		
Labels	**Train**	**Dev**	**Test**	**Train**	**Dev**	**Test**	**Train**	**Dev**	**Test**
Act	5631	651	327	8111	1072	976	9664	1368	1186
Address	1230	147	133	4277	750	2197	9921	1544	2874
Court	727	101	109	1123	209	202	1305	217	205
EffectiveDate	468	83	59	786	148	196	1179	188	236
PII_Ref	41	20	11	270	40	57	445	96	72
Parties	2764	344	262	3747	895	711	5627	1145	833
Percentage	446	42	20	518	89	81	550	94	84
Price	23	1	2	96	19	13	96	22	14
Principal	-	-	-	150	21	24	244	31	47
Ratio	26	4	4	92	16	13	151	21	13
Regulation	1085	144	104	1212	153	143	1484	181	188
RenewalTerm	120	20	8	120	20	8	120	20	8
Rent	-	-	-	-	-	-	32	4	6
Role	1534	135	114	1534	135	114	1756	148	126
Salary	288	33	15	288	33	15	317	35	17
Shares	61	7	7	104	19	13	108	22	16
TerminationDate	233	34	25	250	40	28	302	48	34
Title	1054	92	95	1439	256	183	2293	338	240
O	227962	27836	21804	360225	63879	73294	495693	87075	86334

Table 3: Label-wise data distribution statistics

are more effectively managed using prompts.

4.5 Analysis

The primary focus of our research was placed on evaluating the generalization capabilities of prompt based models in extracting information from unseen and new contract categories. To achieve this, we incorporated paragraphs from recent contracts sourced from SEC EDGAR as part of unseen data, expanding the diversity of contract types beyond the model's training data.

Remarkably, despite not being explicitly trained on *Severance* and *Transition* agreements, the instruction-based T5-small model demonstrated an impressive ability to accurately identify the correct contract titles in these unseen categories. The same can be observed from the output sample presented in Appendix C Figure 8. This indicates promising generalization skills, which are crucial for real-world applications where encountering diverse and evolving contract types is common. Moreover, we observed situations where the training and test samples for certain entities, such as share price and rental amount, were relatively limited. Despite this constraint, the model exhibited strong predictive capabilities, achieving commendable accuracy in extracting these specific entities as shown in Appendix C Figure 7 & 9.

By inducing data in stages and analyzing the inferences on models trained on data from these stages, we conclude that the overall efficiency can be enhanced by increasing a small set of annotated samples as observed in Appendix B Figure 4, 5 & 6. This represents a positive advancement, particu-

	Token Classification (LegalBERT)			RoBERTa-Large + LoRa			T5-small Instruction Model + LoRa		
Entity Name	**precision**	**recall**	**f1-score**	**precision**	**recall**	**f1-score**	**precision**	**recall**	**f1-score**
Act	0.50	0.64	0.56	0.32	0.18	0.23	0.83	**1.00**	0.91
Address	0.35	0.33	0.34	0.19	0.17	0.18	1.00	**0.67**	0.80
Court	0.64	0.70	0.67	0.42	0.44	0.43	1.00	**1.00**	1.00
EffectiveDate	0.71	0.72	0.72	0.62	0.53	0.57	0.94	**0.94**	0.94
PII_Ref	1.00	**1.00**	1.00	0.08	0.06	0.07	1.00	**1.00**	1.00
Parties	0.43	0.64	0.52	0.24	0.18	0.21	1.00	**0.74**	0.85
Percentage	0.81	0.73	0.77	0.72	0.81	0.76	0.78	**1.00**	0.88
Price	0.93	**1.00**	0.96	0.90	0.45	0.60	1.00	0.80	0.89
Principal	0.41	0.44	0.43	0.12	0.12	0.12	0.80	**0.67**	0.73
Ratio	0.50	0.62	0.56	0.30	0.50	0.37	0.25	**0.67**	0.36
Regulation	0.60	**0.88**	0.71	0.21	0.22	0.22	0.66	0.75	0.70
RenewalTerm	0.50	**0.50**	0.50	0.29	0.29	0.29	0.75	0.43	0.54
Rent	-	-	-	-	-	-	0.50	**1.00**	0.67
Role	0.66	0.76	0.70	0.8	0.88	0.83	0.33	**1.00**	0.50
Salary	0.52	**0.88**	0.65	0.24	0.23	0.24	0.67	0.67	0.67
Shares	-	-	-	0.39	0.63	0.48	1.00	**1.00**	1.00
TerminationDate	0.71	**0.92**	0.80	0.44	0.58	0.50	1.00	0.67	0.80
Title	0.68	0.79	0.73	0.37	0.22	0.28	1.00	**0.84**	0.91

Table 4: Model Comparisions on Overall Test Dataset (Stage-3 Dataset).

larly given the challenge of acquiring large quantities of annotated data for all contract titles, which can be arduous and resource-intensive. Our findings highlight the robustness of instruction-based models, emphasizing their potential to adapt and perform well in scenarios with sparse data and novel contract categories. This versatility augments the applicability of such models in real-world settings, where access to exhaustive training data can be challenging. Higher recall is another desirable feature, and we observed that the instruction-based T5-small model trained on the entire dataset achieved a higher recall value compared to other baseline models.

5 Conclusion and Future work

Our contributions delved into the landscape of named entity taggers such as Spacy, legal entity taggers like LexNLP, and few-shot instruction models like ChatGPT for extracting entities from contracts. While these models offered valuable capabilities, we identified certain limitations that we aimed to address in our work.

A significant drawback we observed in most existing models was the absence of fine-grained classification. For instance, currency terms were often tagged as mere amounts, without further distinguishing whether they represented salary, rent, principal amount, or credit amount. Similarly, the classification for dates lacked specificity, leaving out information on whether they referred to effective dates, start dates, or termination dates. Moreover, the zero-shot and few-shot learning of GPT models proved insufficient for accurate predictions in many cases, emphasizing the need for further fine-tuning on task-specific data.

To tackle these limitations, our paper focused on providing fine-grained classification for general entities such as amounts and dates, although we acknowledge that our approach might not cover all possible scenarios exhaustively. Additionally, we undertook the task of fine-tuning instruction-based models and made the resulting dataset and models publicly available.

Looking ahead, we recognize the potential benefits of applying fine-grained classification to other entities, such as percentages, and intend to explore this avenue in our future work. Furthermore, we aspire to widen the scope of contract categories addressed in our research to ensure a more comprehensive and practical solution. Our paper contributes to the advancement of entity extraction from contracts by addressing crucial limitations and providing fine-grained classification. We hope our efforts will inspire further research and improvements in this domain.

Limitations

Despite efforts to collect a broad range of contracts from various sources, our dataset may not fully represent the entire spectrum of contract types and variations. As a result, certain contract entities might be underrepresented or not covered at all, leading to potential biases in the extraction results.

Another important limitation is the unavailability of computational resources, preventing us from evaluating the fine-tuning capabilities of decoder-only models, such as GPT variants and Llama 2-Chat. Due to restricted access to high-performance computing facilities, we adopt a checkpoint-based training approach instead of training on the entire dataset at once. As a consequence, there might be a slight reduction in the performance of our models.

Ethics Statement

In this paper, we utilized a dataset comprising European contracts sourced from SEC EDGAR, which is a publicly available repository. The annotation process for this dataset was carried out by student annotators, who do not possess expertise in the legal field. However, to ensure the accuracy and quality of annotations, the evaluation was conducted by professionals with substantial experience in dealing with contracts on a frequent basis.

Ethical considerations were paramount throughout this research endeavor. All data used in this study were sourced from publicly available and legally accessible repositories, and appropriate attribution and compliance with copyright regulations were maintained. Moreover, the student annotators were provided with clear guidelines to ensure that the annotation process was conducted with precision and fairness.

References

Khalid Al-Kofahi, Alex Tyrrell, Arun Vachher, and Peter Jackson. 2001. A machine learning approach to prior case retrieval. In *Proceedings of the 8th international conference on Artificial intelligence and law*, pages 88–93.

Ting Wai Terence Au, Ingemar J Cox, and Vasileios Lampos. 2022. E-ner–an annotated named entity recognition corpus of legal text. *arXiv preprint arXiv:2212.09306*.

Valentin Barriere and Amaury Fouret. 2019. May i check again?–a simple but efficient way to generate and use contextual dictionaries for named entity recognition. application to french legal texts. *arXiv preprint arXiv:1909.03453*.

Paheli Bhattacharya, Kaustubh Hiware, Subham Rajgaria, Nilay Pochhi, Kripabandhu Ghosh, and Saptarshi Ghosh. 2019. A comparative study of summarization algorithms applied to legal case judgments. In *Advances in Information Retrieval: 41st European Conference on IR Research, ECIR 2019, Cologne, Germany, April 14–18, 2019, Proceedings, Part I 41*, pages 413–428. Springer.

Kalina Bontcheva, Hamish Cunningham, Ian Roberts, Angus Roberts, Valentin Tablan, Niraj Aswani, and Genevieve Gorrell. 2013. Gate teamware: a web-based, collaborative text annotation framework. *Language Resources and Evaluation*, 47:1007–1029.

Ilias Chalkidis and Ion Androutsopoulos. 2017. A deep learning approach to contract element extraction. In *JURIX*, pages 155–164.

Ilias Chalkidis, Ion Androutsopoulos, and Nikolaos Aletras. 2019. Neural legal judgment prediction in english. *arXiv preprint arXiv:1906.02059*.

Ilias Chalkidis, Ion Androutsopoulos, and Achilleas Michos. 2017. Extracting contract elements. In *Proceedings of the 16th edition of the International Conference on Articial Intelligence and Law*, pages 19–28.

Qian Chen, Zhu Zhuo, and Wen Wang. 2019. Bert for joint intent classification and slot filling. *arXiv preprint arXiv:1902.10909*.

Melonie de Almeida, Chamodi Samarawickrama, Nisansa de Silva, Gathika Ratnayaka, and Amal Shehan Perera. 2020. Legal party extraction from legal opinion text with sequence to sequence learning. In *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, pages 143–148.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Erik-Lân Do Dinh, Richard Eckart de Castilho, and Iryna Gurevych. 2015. In-tool learning for selective manual annotation in large corpora. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pages 13–18.

Jenny Rose Finkel, Trond Grenager, and Christopher D Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd annual meeting of the association for computational linguistics (ACL'05)*, pages 363–370.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

Zhiheng Huang and Wei Xu. 2015. Kai yu. *Bidirectional LSTM-CRF models for sequence tagging. CoRR, abs/1508.01991*.

Peter Jackson, Khalid Al-Kofahi, Alex Tyrrell, and Arun Vachher. 2003. Information extraction from case law and retrieval of prior cases. *Artificial Intelligence*, 150(1-2):239–290.

Prathamesh Kalamkar, Astha Agarwal, Aman Tiwari, Smita Gupta, Saurabh Karn, and Vivek Raghavan. 2022. Named entity recognition in indian court judgments. *arXiv preprint arXiv:2211.03442*.

Jan-Christoph Klie, Michael Bugert, Beto Boullosa, Richard Eckart de Castilho, and Iryna Gurevych. 2018. The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In *proceedings of the 27th international conference on computational linguistics: system demonstrations*, pages 5–9.

Jonathan K Kummerfeld. 2019. Slate: a super-lightweight annotation tool for experts. *arXiv preprint arXiv:1907.08236*.

Nikolaos Lagos, Frederique Segond, Stefania Castellani, and Jacki O'Neill. 2010. Event extraction for legal case building and reasoning. In *Intelligent Information Processing V: 6th IFIP TC 12 International Conference, IIP 2010, Manchester, UK, October 13-16, 2010. Proceedings 6*, pages 92–101. Springer.

E Leitner, G Rehm, and J Moreno-Schneider. A dataset of german legal documents for named entity recognition. arxiv 2020. *arXiv preprint arXiv:2003.13016*.

Spyretta Leivaditi, Julien Rossi, and Evangelos Kanoulas. 2020. A benchmark for lease contract review. *arXiv preprint arXiv:2010.10386*.

Bill Yuchen Lin, Dong-Ho Lee, Frank F Xu, Ouyu Lan, and Xiang Ren. 2019. Alpacatag: An active learning-based crowd annotation framework for sequence tagging. In *Proceedings of the 57th Conference of the Association for Computational Linguistics*.

Yupian Lin, Tong Ruan, Ming Liang, Tingting Cai, Wen Du, and Yi Wang. 2022. DoTAT: A domain-oriented text annotation tool. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 1–8, Dublin, Ireland. Association for Computational Linguistics.

Haokun Liu, Derek Tam, Mohammed Muqeeth, Jay Mohta, Tenghao Huang, Mohit Bansal, and Colin A Raffel. 2022. Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. *Advances in Neural Information Processing Systems*, 35:1950–1965.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach.

Xuezhe Ma and Eduard Hovy. 2016. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1064–1074, Berlin, Germany. Association for Computational Linguistics.

Vijit Malik, Rishabh Sanjay, Shubham Kumar Nigam, Kripa Ghosh, Shouvik Kumar Guha, Arnab Bhattacharya, and Ashutosh Modi. 2021. Ildc for cjpe: Indian legal documents corpus for court judgment prediction and explanation. *arXiv preprint arXiv:2105.13562*.

Stephen Mayhew and Dan Roth. 2018. Talen: Tool for annotation of low-resource entities. In *Proceedings of ACL 2018, System Demonstrations*, pages 80–86.

Sudip Mittal, Karuna P Joshi, Claudia Pearce, and Anupam Joshi. 2015. Parallelizing natural language techniques for knowledge extraction from cloud service level agreements. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2831–2833. IEEE.

Hiroki Nakayama, Takahiro Kubo, Junya Kamura, Yasufumi Taniguchi, and Xu Liang. 2018. doccano: Text annotation tool for human. *Software available from https://github. com/doccano/doccano*, page 34.

Minh-Quoc Nghiem and Sophia Ananiadou. 2018. Aplenty: annotation tool for creating high-quality datasets using active and proactive learning. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 108–113.

Joel Niklaus, Veton Matoshi, Pooja Rani, Andrea Galassi, Matthias Stürmer, and Ilias Chalkidis. 2023. Lextreme: A multi-lingual and multi-task benchmark for the legal domain. *arXiv preprint arXiv:2301.13126*.

Sampo Pyysalo, Jorge Campos, Juan Miguel Cejuela, Filip Ginter, Kai Hakala, Chen Li, Pontus Stenetorp, and Lars Juhl Jensen. 2015. Sharing annotations better: Restful open annotation. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pages 91–96.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.

Paulo Silva, Carolina Gonçalves, Carolina Godinho, Nuno Antunes, and Marilia Curado. 2020. Using natural language processing to detect privacy violations in online contracts. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 1305–1307.

Jana Straková, Milan Straka, and Jan Hajic. 2014. Open-source tools for morphology, lemmatization, pos tagging and named entity recognition. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 13–18.

Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, and Nikolai Liubimov. 2020-2022. Label Studio: Data labeling software. Open source software available from https://github.com/heartexlabs/label-studio.

Liwen Wang, Rumei Li, Yang Yan, Yuanmeng Yan, Sirui Wang, Wei Wu, and Weiran Xu. 2022. Instructionner: A multi-task instruction-based generative framework for few-shot ner. *arXiv preprint arXiv:2203.03903*.

Chaojun Xiao, Haoxi Zhong, Zhipeng Guo, Cunchao Tu, Zhiyuan Liu, Maosong Sun, Yansong Feng, Xianpei Han, Zhen Hu, Heng Wang, et al. 2018. Cail2018: A large-scale legal dataset for judgment prediction. *arXiv preprint arXiv:1807.02478*.

Puyang Xu and Ruhi Sarikaya. 2013. Convolutional neural network based triangular crf for joint intent detection and slot filling. In *2013 ieee workshop on automatic speech recognition and understanding*, pages 78–83. IEEE.

Dan Yang, Christina Leber, Luis Tari, Aravind Chandramouli, Andrew Crapo, Richard Messmer, and Steven Gustafson. 2013. A natural language processing and semantic-based system for contract analysis. In *2013 IEEE 25th International Conference on Tools with Artificial Intelligence*, pages 707–712. IEEE.

Seid Muhie Yimam, Chris Biemann, Richard Eckart de Castilho, and Iryna Gurevych. 2014. Automatic annotation suggestions and custom annotation layers in webanno. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 91–96.

Seid Muhie Yimam, Iryna Gurevych, Richard Eckart de Castilho, and Chris Biemann. 2013. Webanno: A flexible, web-based and visually supported system for distributed annotations. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 1–6.

Baoli Zhang, Zhucong Li, Zhen Gan, Yubo Chen, Jing Wan, Kang Liu, Jun Zhao, Shengping Liu, and Yafei Shi. 2021. Croano: A crowd annotation platform for improving label consistency of chinese ner dataset. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 275–282.

Ruixue Zhang, Wei Yang, Luyun Lin, Zhengkai Tu, Yuqing Xie, Zihang Fu, Yuhao Xie, Luchen Tan, Kun Xiong, and Jimmy Lin. 2020. Rapid adaptation of bert for information extraction on domain-specific business documents. *arXiv preprint arXiv:2002.01861*.

A Guidelines

We detail the guidelines followed while annotating and creating the corpus for contract NER in Table 5.

B Predictions from Instruction Model

Figures of predictions from the instruction-based T5-small model are presented here.

C Finetuned T5-small instruction model inferences

Inferences for T5-small instruction model trained on Stage3 data on latest contracts from SEC.

ContractTitle	**Definition:** The name of the document. Clue or helping words ["Agreement", "Note", "Contract"] **Text:** 'htm EX 10.2 2046 NOTE Exhibit EXHIBIT 10.2PROMISSORY NOTE$450,000,000 NEW YORK' **Ans:** PROMISSORY NOTE **Text:** 'This Loan Agreement (the Agreement) is made this 28th day of September, 2017 by and between BRANCH BANKING AND TRUST COMPANY, a North Carolina banking corporation (Bank), and:' **Ans:** Loan Agreement
ContractParties	**Definition:** The two or more parties who signed the contract. Clue or helping words ["parties", "agreement between"] **Text:** 'Arch U.S. MI Holdings Inc., a Delaware corporation (the "Company"), hereby promises to pay the principal amount of four hundred fifty million U.S. dollars ($450,000,000) to Arch Capital Finance LLC, a limited liability company.' **Ans:** Arch U.S. MI Holdings Inc., Arch Capital Finance LLC, a limited liability company
EffectiveDate	**Definition:** On what date is the contract effective? Clue or Helping words ["Effective Date", "entered into"] **Text:** 'TO EXECUTIVE EMPLOYMENT AGREEMENT is entered into as of August 8, 2016, to be effective as of July 1, 2016, by and between Aqua Metals, Inc., a Delaware corporation ("Company"), and Thomas Murphy ("Executive").' **Ans:** July 1, 2016
TerminationDate	**Definition:** On what date will the contract's initial term expire or the maturity Date? Clue or Helping words ["terminates on" , "termination date", "maturity date", "valid till", "closing date"] **Text:** 'at the Holder's office located at 360 Hamilton Avenue, Suite 600, White Plains, New York 10601 or at such other address as the Holder shall direct, on December 1, 2046 (the "Maturity Date") and to pay interest as described below' **Ans:** December 1, 2046
RenewalDate	**Definition:** What is the renewal term after the initial term expires? This includes automatic extensions and unilateral extensions with prior notice. Clue or Helping words ["will be renewed", "revived"] **Text:** 'This Agreement shall begin upon the date of its execution by MA and acceptance in writing by Company and shall remain in effect until the end of the current calendar year and shall be automatically renewed for successive one (1) year periods unless otherwise terminated.' **Ans:** successive one (1) year periods
SalaryCompensation	**Definition:** Salary mentioned for the employee in Employee Agreements. Clue or Helping words ["salary", "annual salary", "compensation", "pay"] **Text:** 'The Employee will be paid an annual salary of Three Hundred Eighty Thousand Dollars ($380,000). ' **Ans:** $380,000
GoverningLaw	**Definition:** Which state/country's law governs the interpretation of the contract? Clue or Helping words ["Governing Law", "fall under", "jurisdiction"] **Text:** '2.7.Applicable Law. This Guaranty shall be construed in accordance with and governed by the laws of the State of New York, without regard to conflict of laws principles.' **Ans:** State of New York
Employment Role	**Definition:** The role for which an employee is employed. Clue or Helping words ["employed", "appointed as"] **Text:** 'Employee serves as its Executive Vice President Chief Commercial Banking Officer responsible for managing and coordinating the Bank's commercial banking activities.' **Ans:** Executive Vice President Chief Commercial Banking Officer
NumberOfShares	**Definition:** The number of shares allocated. Clue or Helping words ["shares", "allocated", "purchased"] **Text:** The Trust desires to sell and the Corporation desires to purchase 1,500,000 shares of common stock, $1.00 par value per share, of the Corporation (the "Stock").' **Ans:** 1,500,000 shares
SharePrice	**Definition:** Price per share.Clue or Helping words ["per value", "per unit", "per share"] **Text:** 'The Trust desires to sell and the Corporation desires to purchase 1,500,000 shares of common stock, $1.00 par value per share, of the Corporation (the "Stock").' **Ans:** $1.00
Act	**Definition:** An Act is a type of legislation that has been passed by a legislative body, such as a parliament or Congress. Clue or Helping words ["act"] **Text:** 'The Executive's covered dependents at the time of termination shall be entitled to all benefits under the Company's welfare benefit plans (within the meaning of Section 3(1) of the Employee Retirement Income Security Act of 1974, as amended).' **Ans:** Employee Retirement Income Security Act of 1974
PII Reference	**Definition:** PII stands for Personally Identifiable Information. This may include information such as name, address, date of birth, Social Security number, and other identifying information. **Text:** 'Atlantic Stewardship Bank630 Godwin AvenueMidland Park, NJ 07432-1405' **Ans:** 07432-1405
Regulation	**Definition:** A regulation is a rule or order that has been issued by an administrative agency or other government body, usually to implement or interpret a law. **Text:** 'The Claims released include any alleged violation by the Company of: Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e et seq. ' **Ans:** 42 U.S.C. § 2000e **Text:** 'The Employment Retirement Income Security Act of 1974, as amended, 29 U.S.C. § 1001 et seq. ' **Ans:** 29 U.S.C. § 1001
Principal Amount	**Definition:** The initial sum of money borrowed or invested, excluding interest and other charges. Clue or Helping words ["principal amount", "initial sum", "borrowed amount", "invested amount"] **Text:** 'The principal amount of the loan is $10,000.' **Ans:** $10,000
Revolving Credit	**Definition:** A type of credit facility that allows a borrower to repeatedly access a specified credit limit as long as the terms of the agreement are met, and the outstanding balance is repaid in full or partially. Clue or Helping words ["revolving credit", "credit facility", "credit limit", "borrowing", "repeated access"] **Text:** 'The borrower has access to a revolving credit line with a limit of $5,000' **Ans:** $5,000
Rental Amount	**Definition:** The specified payment to be made by a tenant to a landlord in exchange for the use and occupancy of a property or asset. Clue or Helping words ["rental amount", "payment", "tenant", "landlord", "monthly rent"] **Text:** 'The monthly rental amount for the apartment is $1,500' **Ans:** $1,500

Table 5: Named Entities with definition and examples

Figure 4: Generalization-Capability over data stages.

Figure 5: Improvement in classification efficiency

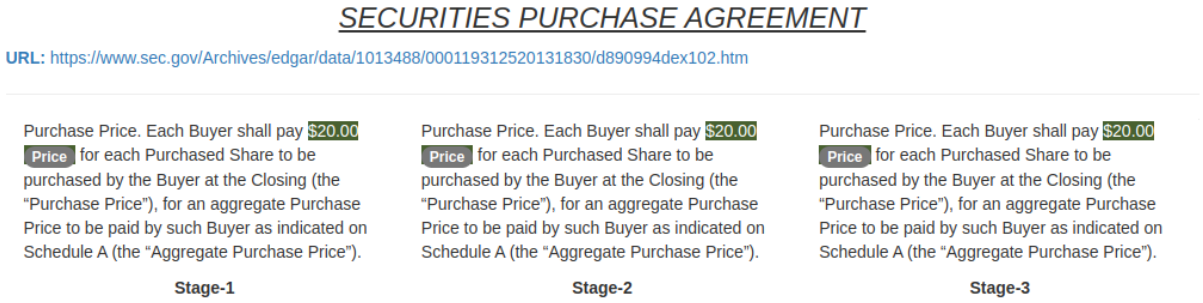

Figure 6: Efficient classification for underrepresented entities

Figure 7: Efficient recognition of underrepresented entity *Rental Amount*

Title: EMPLOYMENT CONTRACT

Url: https://www.sec.gov/Archives/edgar/data/1883814/000182912623001536/legato2_ex10-6ii.htm

Text: During Employee's employment with the Employer, Employee is eligible to receive a one-time, lump sum reimbursement of up to, but not exceeding, Two Thousand Dollars and no Cents ($2,000.00), less applicable and authorized taxes, deductions, withholdings for Employee's reasonable and necessary attorneys' fees incurred in connection with the review, negotiation and execution of this Agreement.

Entities: Two Thousand Dollars is a Principal.

Figure 8: Efficient recognition of underrepresented entity *Principal Amount*

Title: TRANSITION AND RETIREMENT AGREEMENT

Url: https://www.sec.gov/Archives/edgar/data/886835/000119312521271278/d221041dex101.htm

Text: This Transition and Retirement Agreement (this "Transition Agreement") is by and between Superior Energy Services, Inc., a Delaware corporation (the "Company"), and Alan Patrick Bernard (the "Executive"). The parties agree as follows:", "pred_text": "Transition and Retirement Agreement is a Title, Superior Energy Services, Inc., Alan Patrick Bernard is a Parties.

Entities: Transition and Retirement Agreement is a Title, Superior Energy Services, Inc., Alan Patrick Bernard is a Parties.

Figure 9: Entity Recognition of unseen contract titles