Relation Representation Learning for Special Cargo Ontology

Relation Representation Learning for Special Cargo Ontology

Vahideh Reshadat
Department of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
v.reshadat@tue.nl

Alp Akçay
Department of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
a.e.akcay@tue.nl

Kalliopi Zervanou
Department of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
k.zervanou@tue.nl

Yingqian Zhang
Department of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, The Netherlands
yqzhang@tue.nl

Eelco de Jong
Validaide B.V
Amsterdam, The Netherland
eelco.dejong@validaide.com

Abstract—**Non-transparent shipping processes of transporting goods with special handling needs (special cargoes) have resulted in inefficiency in the airfreight industry. Special cargo ontology elicits, structures, and stores domain knowledge and represents the domain concepts and relationship between them in a machine-readable format. In this paper, we proposed an ontology population pipeline for the special cargo domain, and as part of the ontology population task, we investigated how to build an efficient information extraction model from low-resource domains based on available domain data for industry use cases. For this purpose, a model is designed for extracting and classifying instances of different relation types between each concept pair. The model is based on a relation representation learning approach built upon a Hierarchical Attention-based Multi-task architecture in the special cargo domain. The results of experiments show that the model could represent the complex semantic information of the domain, and tasks initialized with these representations achieve promising results.**

Keywords—*Transportation Ontology*, S*pecial Cargo Domain*, *Relation Extraction, Natural Language Processing*

I. INTRODUCTION

The airfreight industry suffers from a lack of knowledge standardization, poor and non-transparent shipping processes [1]. The products and goods with special handling need during transportation such as temperature-sensitive pharmaceuticals, live animals, dangerous goods (e.g., lithium batteries), flowers, and food products are known as special cargo or special freight. A concrete example in healthcare shipping is the distribution of COVID-19 vaccines using temperature-sensitive distribution systems that requires careful route planning supported by industry stakeholders. Routing based on carriers, services, and other conditions is a complex task, as there are many combinations for freight shipments.

Airfreight forwarders (i.e., the individuals or companies that plan the transportation of shipments) need to know all information concerning the routing, such as guidelines, restrictions, and risks. Currently, the forwarders are relying on expert knowledge, and most operations are handled manually. This is difficult due to the complexity of specific product features (e.g., different types of chemical products such as lithium batteries) and the lack of transparency and standardization of capabilities and services offered by the air freight supplier [1] . Manually gathering this information is problematic and time-consuming and affects the efficiency of the process. There might be possible selections that could lead to a lower-cost efficiency. Hence, route planning at freight forwarders has significant potential for optimization and efficiency with the application of advanced data analytics and Artificial Intelligence (AI) [1].

Automating route planning and comparison requires a knowledge base of special cargo services, and domain information elicitation is crucial in developing and populating this special cargo ontology. Due to the lack of domain data, training a model for extracting information and populating the cargo ontology is a challenging task. In this paper, we collect and structure the available knowledge related to shipping and cargo services from a variety of sources to populate the special cargo ontology developed based on knowledge elicitation and UPON [2] methodology [2], [3] in LARA project [4], [5]. Information Extraction (IE), as a part of the ontology population task, is the process of extracting relevant information from massive data, such as corpora, web, database, semi-structured, and multimedia documents that has a wide range of applications in Natural Language Processing (NLP) tasks such as question answering [6], [7] text classification [8], [9] semantic similarity [10], etc.

---

[1] Air carriers and Ground Handling Agents (GHA)

[2] Unified Process for Ontology

Essential information in cargo services can be found on the web pages of airline and airport companies. However, each company provides the information in a different format, using textual descriptions, tables, and possibly different languages, etc. The terms and concepts used in the descriptions are not standardized and contain a high diversity and terminology difference. The lack of terminology resources for the cargo domain is another challenge and due to the large amount of data and maintenance problems, manual extraction of such information is costly.

This research focuses on the automatic information extraction from cargo-related texts using deep neural networks. The scarcity of the annotated domain data makes the training of a robust classifier difficult. This research sheds light on the problem of eliciting special cargo information for populating the domain ontology given minimum human involvement and in the absence of enough training data for target task. We first develop a model by automatically generated domain-specific train data for simple tasks in a hierarchical multi-task fashion; then, we use this pre-trained model as a feature extractor for the target task. This paper makes the following contributions:

- We propose a special cargo ontology population pipeline by using deep neural networks to extract information from the domain documents.

- A relation representation model is proposed by a hierarchical attention-based multi-task architecture that achieves reasonable performance with limited domain-specific training data.

- We provide some practical lessons on developing and designing the deep neural relation extractors in the special cargo shipment domain w.r.t. the balance between domain data and efficiency.

The rest of the paper is organized as follows: In Section II, related works are reviewed; Section III presents our proposed model in detail; Section IV and Section V are dedicated to the details of the datasets, experiments and results; and finally, Section VI concludes the paper.

II.   RELATED WORK

Numerous relation extraction methods for the ontology population task have been proposed in recent years that support a large variety of different approaches, from simple rule-based [11] and statistical methods [12] to complex machine learning [13] and hybrid structures [14]. Rule-based methods are based on some predefined rules that describe the structure of the required information. Hearst [11] proposed a well-known domain-independent rule-based method that is based on the lexico-syntactic patterns. These patterns are generated by using the bootstrapping of a set of seed instances. The method can be

used for constructing the taxonomic relations in the ontology by extracting hyponymy relations from large corpora. Rule-based methods need a comprehensive understanding of the domain. Massive demand for human intervention is the main drawback of these methods [15]. Machine learning approaches have been developed to overcome these limitations. They are the most widely used methods in ontology population and are divided into three main categories, namely supervised [16], weakly supervised [17], and unsupervised [18] learning methods. Generating annotated text for supervised methods is costly and time-consuming. Semi-supervised and unsupervised approaches alleviate this problem by using less or no training data.

Since the fundamental distinguishing characteristic of deep neural network models w.r.t other machine learning methods is that the feature extraction process is done automatically, deep neural networks are the state-of-the-art methods that have attracted increasing interest in the ontology population task. Some methods consider relation classification and exploit different deep neural network models such as convolutional deep neural networks [19], [20] to extract lexical and context level features from sentences. These features are then fed into a softmax classifier to predict the relation type between a pair of name entities. These methods often need a tagged corpus in which the concepts and the semantic relations among them are annotated. Recent state-of-the-art solutions rely on using pre-trained models to achieve high performance [21]. A task-independent relation representation method is proposed in [22] that builds representations based on the entity-linked text. This method was inspired by Harris' distributional hypothesis and the recent advances in text representation learning, specifically BERT [23]. These models are designed for the general text domains, and few works were done on relation extraction for ontology learning in specific domains.

III.   METHODOLOGY

In this section, we introduce the pipeline of the special cargo ontology population. Since the focus of the paper is on the relation representation task, we describe the methodology used in the Relation Extractor component with more details.

Due to the lack of domain training data, developing and populating special cargo ontology is challenging. We generated specific conceptualizing of special cargo shipment based on UPON methodology [2], [3]. This preliminary work structures the available domain knowledge with elicitation techniques to derive knowledge from domain experts [3]. In order to determine a choice set for the routing options of special cargo shipment, concepts need to be instantiated with domain data.

As shown in Fig. 1, in the Preprocessing step, online documents are processed to remove images, tables, references. Instances of concepts and relations are extracted based on the
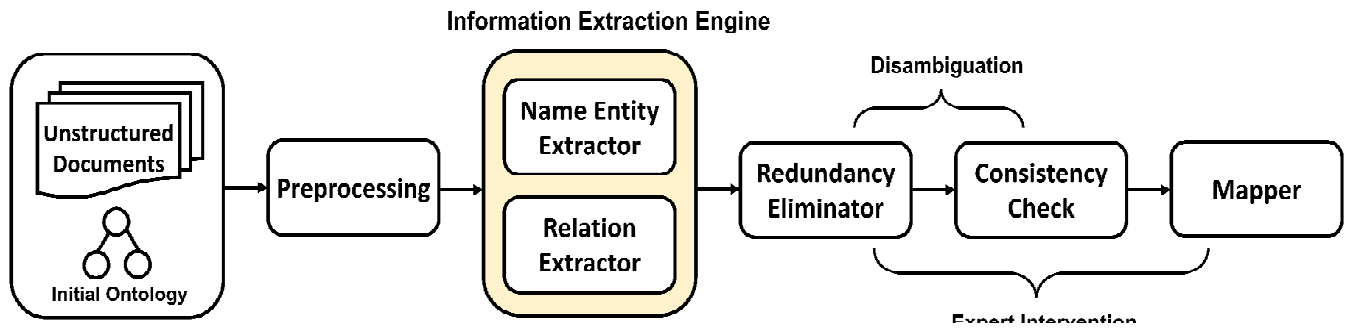
seed ontology in Information Extraction Engine. Redundancy Eliminator eliminates redundant instances of extracted entities and relations. Thereafter, the instances that contradict the knowledge of special cargo ontology are removed. Disambiguation is necessary for performing consistency and redundancy checks. Finally, the result instances are inserted into the special cargo ontology by Mapper. The last three components require expert intervention in providing a set of heuristics and rules. A part of the special cargo ontology is shown in Fig. 2.

There is a massive amount of useful information available on cargo websites, online documents, and databases. For populating this specific ontology, the relevant information should be extracted from these resources and connected coherently in the ontology. Since the existing information extractors are highly dependent on the underlying ontology or knowledge graph used in their design, there is a large variance between the available information extraction systems, which makes applying them difficult in specific domains. Due to the scarcity of labeled training data for the transportation domain, developing an efficient relation extractor for special cargo shipment is difficult and requires significant human investment.

To address this problem, a relation representation model is proposed based on a hierarchical attention-based multi-task architecture. The model is trained using simple tasks where automatically generating domain-specific train data is relatively easy. Then it is applied for initializing a supervised multi-class domain-specific relation classifier with small train data. More details about the model are described in the next section.

A. Learning Relation Representation for Special Cargo Domain Using Hierarchical Multi-task Architecture

Recent language representation models are mostly designed to encode the various sets of linguistic, syntactic, and semantic features for the down-stream tasks [24]. Since multi-task systems take the benefit of inductive transfer among different tasks, they achieve complementary aspects in the encoded representation and, therefore a better generalization performance [25], [26]. Furthermore, training such a model by simple tasks with automatically generated domain-specific training data can produce a rich representation. This inspired us to develop a novel hierarchical model with a combination of three different attention-based NLP models that embeds simple tasks in the low levels of the hierarchy and more complex tasks in the high-level of the hierarchy. Fig. 3 depicts the framework of the model architecture. This model is built upon the previous work [26] where a supervised hierarchical model proposed based on a multi-task architecture.
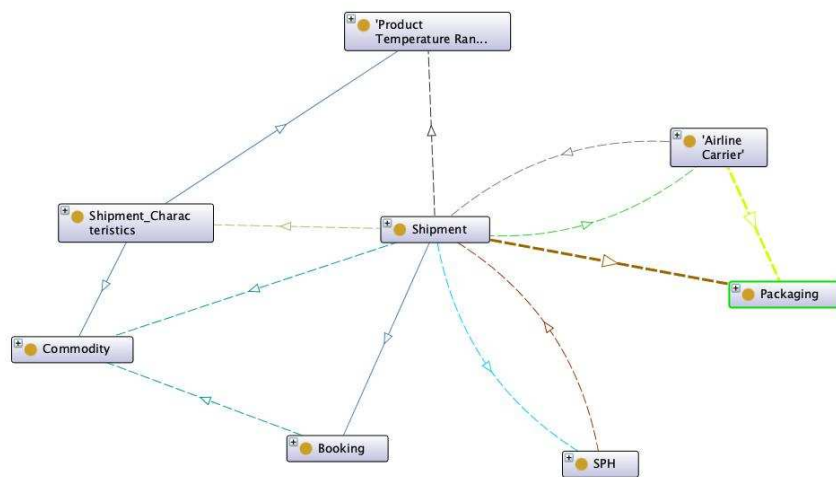

Fig. 2. "Shipment" concept in the spcial cargo ontology

Word Representation: The input of the model is an embedded vector of the input sentence using three different pre-trained models, namely GloVe [27] for word embedding, ElMo [28] for contextual embedding, and a convolutional neural network (CNN) based model [29] for character-level embedding. Therefore, the model takes a sentence $s = (w_1, w_2, ..., w_n)$ as input and encodes word w_t as a concatenation of three pre-trained word embeddings denoted as g_e that provides a rich representation range from character to context level. In order to have a fair comparison with state-of-the-art language representation models (e.g., BERT), such representation models are not applied in the input.

Name Entity Recognition (NER): The first underlying task in the hierarchical model is NER that has a Conditional Random Field (CRF) for detecting NER tags. NER recognizes and classifies name entity mentions in the input sequence. A 2-layer BiLSTM with attention mechanism is used for encoding the input. It takes the word embeddings g_e and outputs sequence embeddings g_{ner} that then fed into CRF-based sequence tagging layer.

Entity Detection (ED): This task is similar to NER but more general in terms of detecting all mentions related to a real life entity, while NER only relies on the name entities. ED is considered as a sequence tagging task that employs a 2-layer BiLSTM with attention mechanism followed by a CRF layer. Thus, the concatenated embedding vector $[g_e, g_{ner}]$ from lower layers is fed into the encoder that outputs embeddings denoted by g_{ed}.

Binary Relation Extraction: The task of identifying semantic relation between entities in the text is Relation Extraction (RE). It requires mention detection and classification of the relation between identified mentions. We adopted a joint model proposed in [30] that jointly learns these subtasks. This binary relation classifier is able to detect whether or not there is a relation between entities in the input sentence.

Having the similar 2-layer BiLSTM with attention on top of the NER and ED tasks, RE encoder takes $[g_e, g_{ed}]$ as input and outputs an embedding denoted g_{re}. These embeddings are used as input to a feed forward network.

There is no apparent agreement about training a hierarchical multi-task model. We applied the effective training method proposed in [31], [26]. For training the model, after each parameter update, a task and a batch of its training data are sampled randomly, and the task is trained. Each task is sampled uniformly, and the training process is iterated until convergence.

Multi-class Relation Extraction: The hierarchical multi-task architecture is trained using domain-specific data. We use this base model as a feature extractor for a new model. The last layer of this binary classification is removed, and the remaining layers are used as a feature extractor for the multi-class relation classifier. Due to the lack of data for training a multi-class relation extraction classifier in the special cargo domain, applying pre-trained models can bring the shared features to the target task. In this case, training the new model using the pre-
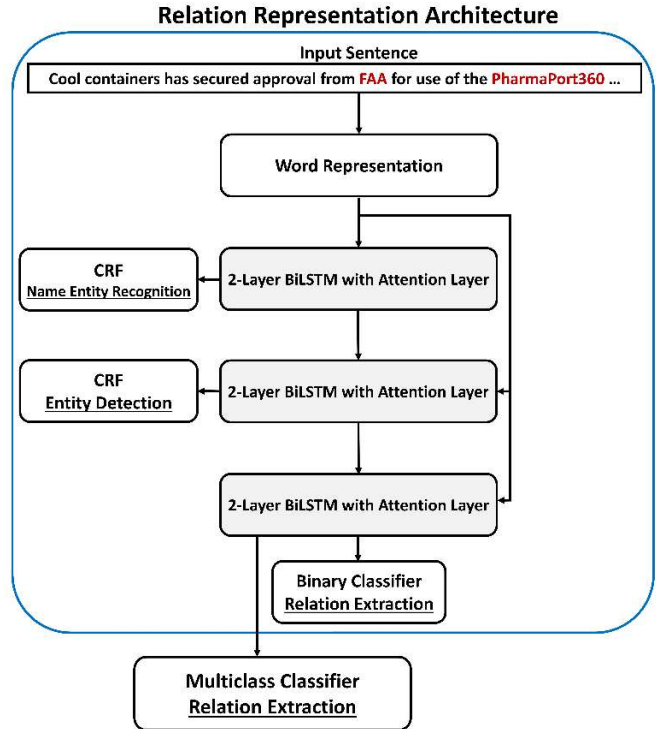

Fig. 3. Architecture of the attention-based hierarchical multi-task relation representation model for multi-class relation classification.

trained model with few samples of target task results in a reasonable performance boost. In other words, a binary relation classifier is trained using the enriched representation in a hierarchical multi-task architecture with a generated corpus of sufficient instances in the cargo domain. The final architecture for the model is obtained, by leveraging the relation representation derived from the hierarchical multi-task architecture as a feature extractor for the multi-class relation extraction task. To the best of our knowledge, no previous research investigates applying an attention-based hierarchical multi-task transforming model as relation representation for multi-class relation classification.

IV. DATASETS

For training and testing different parts of the proposed architectures, different datasets are prepared. These datasets are described in more detail in this section.

A. Train/Test Data for the Attention-based Hierarchical Multi-task Model

In order to train the binary relation classifier, we scraped 28,809 news texts from the cargo news websites that contain formal texts. Since not all articles are related to the special cargo domain, an automatic filtering method is exploited. For this purpose, we used Latent Dirichlet Allocation (LDA) [32] topic model to produce 10 clusters. In contrast to keyword-based methods, topic modeling approaches do not require a predefined set of keywords. The most relevant topic is selected based on the topic containing the most representative terms. To ensure that
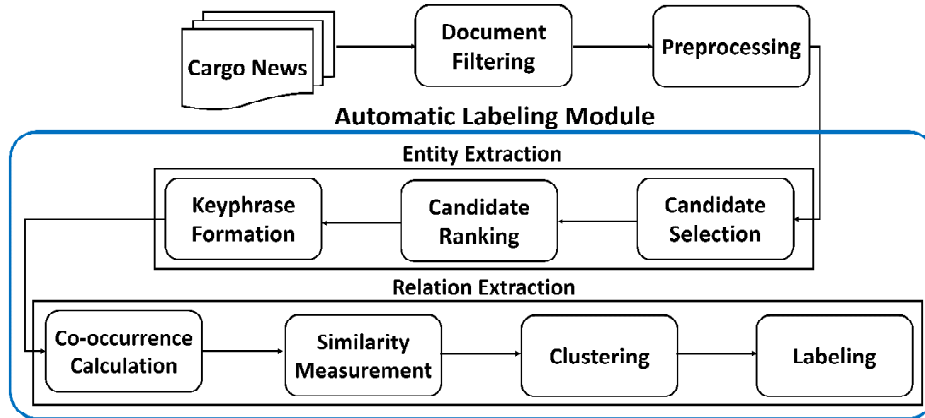
Fig. 4. The workflow of generating dataset for training the hierarchical multi-task model.

news with potentially high relevance is collected, a threshold of 0.9 is applied. Finally, 775 filtered documents are collected.

Various NLP tasks are considered in the preprocessing step that consists of sentence splitting, tokenization, NER, and Part-of-Speech (POS) tagging. Since deep learning models require a huge amount of training data and manually generating this data is expensive, we used a fully automatic labeling policy for annotating train data for special cargo documents. The automatic labeling process consists of two tasks, namely entity extraction and, relation extraction. An unsupervised domain-specific entity extraction framework [33] and a NER tool [34] are used to extract entities from the documents. Fig. 4 shows the different steps for generating the training data.

The goal of automatic labeling is to produce annotations for special cargo news articles with minimum human involvement. There are three main steps for the entity extraction task. The first step is selecting the candidate keywords using heuristics with a list of acceptable POS. In the second step, the selected lexical units are ranked based on the text representation. The ranked candidate keywords are purified, and the final keyphrases are formed in the last step. We used pke [35], an open source python-based keyphrase extraction toolkit that consists of various statistical and graph-based approaches. Based on our evaluations, KPM [36] a statistical-based and, PositionRank [37], a graph-based approach have a higher performance than the other available algorithms for special cargo keyphrase extraction.

A generic clustering-based relation extraction approach consists of four major steps, as shown in Fig. 4. All sentences along with tagged entities from the Entity Extraction task are fed into the Relation Extraction task as input. A pair of name entities and the context between them that occurs within a fixed window size are elicited in a co-occurrence calculation step. The similarity between the contexts of entities that co-occur is measured. There is a wide variety of methods for measuring the similarity between contexts. Levenshtein [38] is a well-known fast lexical similarity measurement that computes the minimum edit distance between two strings. Calculating co-occurrence and similarity are crucial for the context-clustering task. We exploited DBSCAN [39], a density-based clustering method that

doesn't require a prior determination of the number of clusters. Since the combination of KPM and DBSCAN has the highest performance, KPM is selected as a keyphrase extraction method. Labeling is the final step that determines which relation clusters are relevant to the special cargo domain. Patterns from the manually annotated development set are used for labeling the clusters as relevant or irrelevant.

For generating test data, we randomly sampled 223 documents to be annotated manually as relevant or irrelevant for the binary classifier from the filtered documents. Due to irrelevance, 118 documents are removed, and the remaining documents are used in the Dev and Test set. Along with a dataset of 10 domain, online documents are randomly selected. The statistics of the data are depicted in Table I. The training set tends to have more sentences per document than the development set and test set 1, despite all of them being news articles. On the other hand, test set 2 (online documents) has even higher sentences and entities per document, while the number of words per sentence is lower than the training set. This comparison indicates a higher density of relevant information in the online documents compared to the news articles.

TABLE I. STATISTICS OF THE DATASET FOR THE ATTENTION-BASED HIERARCHICAL MULTI-TASK MODEL.

	Train Set	Dev Set	Test Set1	Test Set2
Total documents	552	53	52	10
Total sentences	8361	556	548	221
Sentences per document	15.15	10.49	10.54	22.10

B. Train/Test Data for the Multi-Class Relation Classification Model

There are 43 different relation types in the cargo ontology. Each of these relation types can have two different argument orders. A list of the relation types is selected based on the broad coverage of the relations that participated in special cargo domain ontology [3]. An example of the relation type and a sample sentence for it are shown in Table II. The dataset is provided in the standard format of SemEval-2 Task8 [40] and is available in https://github.com/VahidehReshadat/CargoRelationExtraction.

Available datasets often contain a hundred examples for each relation type. Building such a dataset for the special cargo shipment domain is expensive. For each relation type, only a few instance sentences are selected and labeled. Each sentence is labeled by two domain experts independently. The annotators reached an agreement on 87% of the samples. A subset of the data that annotators have an agreement with is used in the experiments to evaluate multi-class classifiers. The statistics of the datasets are depicted in Table III.

SCR_S contains only a few samples for each relation class in the dataset. SCR_L is created by increasing the number of each class sample in SCR_S. Since some of the relations are similar (e.g., *Arranges* and *Executes*), we created two other datasets, namely SCR_{SM} and SCR_{LM} by merging the similar relations in SCR_S and SCR_L, respectively. These datasets have the least semantic overlap between the relation classes and are coarse-grain representatives of the domain relations.

TABLE II. EXAMPLES OF THE RELATION TYPES AND SAMPLE SENTENCES.

Sentence	Relation Type
<e₁>Brussels Airport</e₁> signs MoU for the seamless transportation of <e₂>COVID-19 vaccine</e₂>.	Ships(e₁,e₂)
<e₁>Active Pharma Ingredients</e₁> is required to be pre-cooled within the desired temperature between <e₂>-25°C and -15°C</e₂> before loading by the shipper.	HasTemperature Range(e₁,e₂)
The <e₁>Va-Q-Tec containers</e₁> keep the <e₂>Pfizer/BioNTech vaccines</e₂> cool enough.	isPackedIn(e₂,e₁)

TABLE III. STATISTICS OF THE GENERATED DATASET FOR THE MULTI-CLASS RELATION CLASSIFIER.

		#annotated samples	#relation types
SCR_S	Train Set	230	43
	Test Set	104	43
SCR_L	Train Set	420	43
	Test Set	104	43
SCR_{SM}	Train Set	180	19
	Test Set	80	19
SCR_{LM}	Train Set	310	19
	Test Set	80	19

V. EXPERIMENTS

In this section, the performance of the special cargo multi-class relation classifiers built on the hierarchical multi-task model is evaluated, and the results are compared with the BERT-base classifier. We apply the BERT relation representation method in [22] for encoding the relations between entity pairs. In this method, the beginning and end of the target entities are marked with special entity markers. Then, the marked input text is fed into the BERT model, and the corresponding states of the beginning of the two entity markers are concatenated, and the relation representation is extracted.

A. Evaluation of the Special Cargo Multi-Class Relation Classifier

In this section, we evaluated the multi-class relation classifier performance over the embedding learned from the hierarchical multi-task model on SCR datasets. Table IV shows the hyper-parameters of the model used in the experiments. In

order to have a fair comparison, the same configuration is applied to the various experiments.

The hierarchical multi-task embedding model can provide an enriched feature resource for the multi-class relation classifier when used for knowledge transfer in the special cargo domain. Table V illustrates the results of the evaluation of the classifier trained using extracted features from the base model. The results are compared with the BERT-based relation classifier in which the BERT-Base model is used as relation representation in the classification task. In this case, the concatenation of the final hidden states corresponding to the entities' start tokens (<e₁> and <e₂>) is considered as the relation representation vector.

The ability of the model to distinguish both fine-grain and coarse-grain relations is promising. The classifier trained using transferred features achieves reasonable results, although not as impressive as that of the BERT-base model; since it doesn't use a big dataset as BERT.

We conduct another set of experiments to measure the performance of the model over relations without any direction. In this case, the instances predicted in the correct class but with a different order of arguments are counted as correct in the performance calculation. The results are shown in Table VI. In summary, the evaluations revealed that the model achieves a higher f-measure across different datasets compared to the directional case. One possible reason is that regardless of direction, only identifying the correct class increases the performance.

TABLE IV. THE HYPER-PARAMETER SETTINGS USED IN THE HIERARCHICAL MULTI-TASK RELATION REPRESENTATION MODEL.

Epochs	20
Learning Rate	0.0005
Dropout Rate (Embedding)	0.5
Dropout Rate (LSTM)	0.2
Batch Size	32
Word Embedding Dimension	100
Char Embedding Dimension	16

TABLE V. EVALUATION RESULTS (F1) FOR THE MULTI-CLASS RELATION CLASSIFICATION TASK.

	$SCRE_{HMTL}$	BERT
SCR_S	43.31	49.5
SCR_L	46.25	52.88
SCR_{SM}	55.73	62.5
SCR_{LM}	56.48	67.09

TABLE VI. EVALUATION RESULTS (F1) FOR MULTI-CLASS RELATION CLASSIFICATION OVER THE NON-DIRECTIONAL DATASETS.

	$SCRE_{HMTL}$	BERT
SCR_S	46.72	53.39
SCR_L	47.18	55.07
SCR_{SM}	59.26	65.38
SCR_{LM}	60.88	71.05

B. Limitations and Practical Implications

In this paper, we discuss the task of extracting special cargo domain information from a variety of text sources. This work sheds more light on the design and development of logistic knowledgebases and the methodology for eliciting domain-specific information. Subsequently, this novel palate of data analytics approach provides a significant role in the freight forwarding industry with a set of solutions for several organizational issues, such as determination of available choice

set for the routing options of shipments with special handling needs. Determination of available routing options is currently very difficult for shippers and forwarders due to the complexity of specific product features (e.g., different types of chemical products such as lithium batteries) and the lack of standardization of capabilities and services offered by airfreight suppliers[3]. Therefore, in order to digitize the manual forwarding process of special cargo, which is inefficient and time-consuming, and to optimize route options, logistics and cargo knowledge needs to be acquired and structured.

We demonstrated that populating industrial scale special cargo ontology from free domain text with automatic labeling and minimum human involvement is practicable. We showed that how applying the proposed representation learning model with only a few samples for each relation type in the dataset can lead to an efficient relation classification model.

Special cargo relation extractor has achieved reasonable performance over the various experiments. For all experiments, the effect of increasing the training data size in the performance is not significant, and applying pre-trained models for learning representation in special cargo domain solely increases the performance of the model and results in an efficient solution for the domain-specific tasks with decreasing the human effort in a large extend.

Our approach still needs to be further improved. Relation extractor built on BERT-based relation representation achieves better results compared to the hierarchical multi-task representation model. This is mainly because of the size of the dataset used for training BERT. While the hierarchical multi-task model solely relies on a relatively small dataset of domain-specific texts, BERT utilizes a gigantic dataset with both general and domain-specific data. Although multi-task structure can take the complementary representation of the relation extraction due to the ordered encoding of the underlying tasks, the size of train data is still effective, and this makes it difficult for SCR$_{EHMTL}$ to reach significant results. Therefore, it is possible to lead to even more improved performance over BERT by increasing the size of training data in the hierarchical multi-task model or utilizing powerful external knowledge and embedding models such as BERT and GPT-3 [41] in its architecture (e.g., in the Word Representation component).

VI. CONCLUSION

In this paper, and as a part of the LARA project, we aim to bring innovation in order to find a robust approach to elicit and model relevant information in the airfreight industry of shipping goods with special handling needs, which requires an urgent need for digital transformation.

Due to the lack of data, training an efficient model is difficult. The severity of this problem intensifies when we are dealing with a specific domain. In this paper, we presented an overview of the special cargo ontology population framework and proposed a relation extractor for the special cargo domain based on a relation representation learning model with attention-based hierarchical multi-task architecture. The hierarchical model is trained on a set of semantic tasks from shallow at the bottom to deep at the top. The input of each task is the outputs of the previous lower tasks in a hierarchical structure. The generated embeddings are then coupled in a new model, and the final classifier is trained. We created different datasets in the special cargo domain for training each of these models. The final model is applied for identifying related associations in populating the domain ontology.

Considering the fact that training the relation representation model needs minimum human intervention, the proposed approach is particularly effective in low-resource regimes and can reduce the human manual effort for creating relation extraction datasets in specific domains. Moreover, it is simple and does not require any external resources to train. To the best of our knowledge, this is the first study that considers information extraction in the cargo transportation domain, and this is the first work in the multi-class relation extraction in the special cargo domain. In future work, we plan to employ external knowledge along with our relation representation. Besides, we will investigate the performance of the attention-based multi-task model in the other tasks of the domain, such as co-reference resolution.

REFERENCES

[1] W. Kersten, T. Blecker, and C. M. Ringle, Digitalization in Supply Chain Management and Logistics: Smart and Digital Solutions for an Industry 4.0 Environment: Berlin: epubli GmbH, 2017.

[2] A. De Nicola, M. Missikoff, and R. Navigli, "A proposal for a unified process for ontology building: UPON," in *International Conference on Database and Expert Systems Applications*, 2005, pp. 655-664.

[3] T. Kolkman, "An Ontology for Special Cargo," *Master's thesis, Eindhoven University of Technology, The Netherlands, Retrieved from https://pure.tue.nl/ws/portalfiles/portal/148941906/Master_Thesis_Tess_Kolkman_.pdf,* 2020.

[4] R. v. d. Mei, "Lane Analysis & Route Advisor (LARA)," *TKI Dinalog,* pp. https://www.dinalog.nl/project/lane-analysis-route-advisor-lara/, 2019.

[5] V. Reshadat, T. Kolkman, K. Zervanou, Y. Zhang, and A. Akcay, "Knowledge Modelling and Incident Analysis for Special Cargo," in *Springer*, ed, 2021.

[6] Y. Pınar, İ. Şahin, and E. Adalı, "Knowledge-Based Question and Answering System for Turkish," *2019 4th International Conference on Computer Science and Engineering (UBMK),* pp. 307-312, 2019.

[7] İ. Dönmez and E. Adalı, "Turkish question answering application with course-grained semantic matrix representation of sentences," in *2017 International Conference on Computer Science and Engineering (UBMK)*, 2017, pp. 6-11.

[8] V. Reshadat and H. Faili, "A New Open Information Extraction System Using Sentence Difficulty Estimation," *Computing and Informatics,* vol. 38, pp. 986–1008, 2019.

[9] V. Reshadat, M. Hourali, and H. Faili, "Confidence Measure Estimation for Open Information Extraction," *Information Systems & Telecommunication,* p. 1, 2018.

[10] V. Reshadat and M.-R. Feizi-Derakhshi, "Studying of semantic similarity methods in ontology," *Research Journal of Applied Sciences, Engineering and Technology,* vol. 4, pp. 1815-1821, 2012.

---

³ Air carriers and Ground Handling Agents (GHA)

[11] M. A. Hearst, "Automatic acquisition of hyponyms from large text corpora," in *Coling 1992 volume 2: The 15th international conference on computational linguistics*, 1992.

[12] D. Maynard, Y. Li, and W. Peters, "NLP Techniques for Term Extraction and Ontology Population," ed, 2008.

[13] M. Craven, D. DiPasquo, D. Freitag, A. McCallum, T. Mitchell, K. Nigam, and S. Slattery, "Learning to construct knowledge bases from the World Wide Web," *Artificial intelligence,* vol. 118, pp. 69-113, 2000.

[14] V. RESHADAT, M. HOORALI, and H. FAILI, "A Hybrid Method for Open Information Extraction Based on Shallow and Deep Linguistic Analysis," *Interdisciplinary Information Sciences,* vol. 22, pp. 87-100, 2016.

[15] M. Cui, L. Li, Z. Wang, and M. You, "A survey on relation extraction," in *China Conference on Knowledge Graph and Semantic Computing*, 2017, pp. 50-58.

[16] [16] A. Fader, S. Soderland, and O. Etzioni, "Identifying relations for open information extraction," in *Proceedings of the 2011 conference on empirical methods in natural language processing*, 2011, pp. 1535-1545.

[17] C. Brewster, F. Ciravegna, and Y. Wilks, "User-centred ontology learning for knowledge management," in *International Conference on Application of Natural Language to Information Systems*, 2002, pp. 203-207.

[18] B. Rosenfeld and R. Feldman, "Clustering for unsupervised relation identification," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, 2007, pp. 411-418.

[19] X. Han, H. Zhu, P. Yu, Z. Wang, Y. Yao, Z. Liu, and M. Sun, "Fewrel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing,* vol. 4803–4809, 2018.

[20] T. Gao, X. Han, H. Zhu, Z. Liu, P. Li, M. Sun, and J. Zhou, "FewRel 2.0: Towards more challenging few-shot relation classification," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing,* pp. 6250–6255, 2019.

[21] C. Li and Y. Tian, "Downstream Model Design of Pre-trained Language Model for Relation Extraction Task," *Association for Computational Linguistics,* pp. 1-15, 2020.

[22] L. B. Soares, N. FitzGerald, J. Ling, and T. Kwiatkowski, "Matching the blanks: Distributional similarity for relation learning," *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics,* pp. 2895–2905, 2019.

[23] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* vol. 1, pp. 4171–4186, 2018.

[24] A. Conneau, D. Kiela, H. Schwenk, L. Barrault, and A. Bordes, "Supervised learning of universal sentence representations from natural language inference data," *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing,* vol. Association for Computational Linguistics, pp. 670-680, 2017.

[25] R. Caruana, "Multitask Learning: A Knowledge-Based Source of Inductive Bias ICML," *Google Scholar Google Scholar Digital Library Digital Library,* 1993.

[26] V. Sanh, T. Wolf, and S. Ruder, "A hierarchical multi-task approach for learning embeddings from semantic tasks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, pp. 6949-6956.

[27] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532-1543.

[28] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* pp. 2227–2237, 2018.

[29] J. P. Chiu and E. Nichols, "Named entity recognition with bidirectional LSTM-CNNs," *Transactions of the Association for Computational Linguistics,* vol. 4, pp. 357-370, 2016.

[30] G. Bekoulis, J. Deleu, T. Demeester, and C. Develder, "Joint entity recognition and relation extraction as a multi-head selection problem," *Expert Systems with Applications,* vol. 114, pp. 34-45, 2018.

[31] A. Søgaard and Y. Goldberg, "Deep multi-task learning with low level tasks supervised at lower layers," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2016, pp. 231-235.

[32] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research,* vol. 3, pp. 993-1022, 2003.

[33] K. S. Hasan and V. Ng, "Conundrums in unsupervised keyphrase extraction: making sense of the state-of-the-art," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, 2010, pp. 365-373.

[34] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. D. Manning, "Stanza: A python natural language processing toolkit for many human languages," *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations,* pp. 101–108, 2020.

[35] F. Boudin, "PKE: an open source python-based keyphrase extraction toolkit," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, 2016, pp. 69-73.

[36] S. R. El-Beltagy and A. Rafea, "KP-Miner: A keyphrase extraction system for English and Arabic documents," *Information systems,* vol. 34, pp. 132-144, 2009.

[37] C. Florescu and C. Caragea, "Positionrank: An unsupervised approach to keyphrase extraction from scholarly documents," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2017, pp. 1105-1115.

[38] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," in *Soviet physics doklady*, 1966, pp. 707-710.

[39] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Kdd*, 1996, pp. 226-231.

[40] I. Hendrickx, S. N. Kim, Z. Kozareva, P. Nakov, D. O. Séaghdha, S. Padó, M. Pennacchiotti, L. Romano, and S. Szpakowicz, "Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals," *Proceedings of the 5th International Workshop on Semantic Evaluation,* pp. 33–38, 2019.

[41] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, and A. Askell, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165,* 2020.