

## A Study on Construction Method of AI based Situation Analysis Dataset for Battlefield Awareness

Yukyung Shin\*, Soyeon Jin\*, Jongchul Ahn\*

\*Engineer, Intelligent C4I Team, Hanwha Systems Co., Seongnam, Korea

\*Chief Engineer, Intelligent C4I Team, Hanwha Systems Co., Seongnam, Korea

\*Chief Engineer, Intelligent C4I Team, Hanwha Systems Co., Seongnam, Korea

### [Abstract]

The AI based intelligent command and control system can automatically analyzes the properties of intricate battlefield information and tactical data. In addition, commanders can receive situation analysis results and battlefield awareness through the system to support decision-making. It is necessary to build a battlefield situation analysis dataset similar to the actual battlefield situation for learning AI in order to provide decision-making support to commanders. In this paper, we explain the next step of the dataset construction method of the existing previous research, 'A Virtual Battlefield Situation Dataset Generation for Battlefield Analysis based on Artificial Intelligence'. We proposed a method to build the dataset required for the final battlefield situation analysis results to support the commander's decision-making and recognize the future battlefield. We developed 'Dataset Generator SW', a software tool to build a learning dataset for battlefield situation analysis, and used the SW tool to perform data labeling. The constructed dataset was input into the Siamese Network model. Then, the output results were inferred to verify the dataset construction method using a post-processing ranking algorithm.

▶ **Key words:** Battlefield Analysis, Battlefield Awareness, Intelligent Command Control System, Artificial Intelligence, Ranking Algorithm

### [요 약]

인공지능에 기반한 지능형 지휘통제체계는 복잡하고 방대한 전장정보와 전술 데이터들을 학습모델을 통해 자동으로 융합 및 추출하여 전장상황을 분석한다. 지휘관은 지능형 지휘통제체계의 상황분석 결과를 제공받아 전장인식이 가능하여 의사결정을 지원할 수 있다. 의사결정지원에 특화된 결과를 지휘관에게 제공하기 위해서는 인공지능을 학습하기 위한 실 전장상황과 유사한 전장상황분석 데이터 셋 생성이 필요하다. 본 논문은 기존 선행연구인 '인공지능 기반 전장상황분석을 위한 가상 전장상황 데이터 셋 생성 연구'의 다음 단계의 데이터셋 구축 방법 연구로 지휘관의 의사결정지원 및 미래 전장인식을 위해 최종적인 전장상황분석 결과에 필요한 데이터셋을 생성하는 방안에 대해 제안하였다. 전장상황 분석용 학습 데이터셋 생성도구 SW를 설계 및 구현하였고, 구현한 SW를 이용하여 데이터 레이블 작업을 진행하였다. Siamese Network 학습모델을 이용하여 구축한 데이터셋을 입력하고, 후처리 알고리즘을 활용한 출력 결과를 도출하여 생성한 데이터셋을 검증하였다.

▶ **주제어:** 전장상황분석, 전장상황인식, 지휘통제체계, 인공지능, 순위 알고리즘

- First Author: Yukyung Shin, Corresponding Author: Yukyung Shin
- \*Yukyung Shin (ykshin@hanwha.com), Intelligent C4I Team, Hanwha Systems Co.
- \*Soyeon Jin (soyeon.jin@hanwha.com), Intelligent C4I Team, Hanwha Systems Co.
- \*Jongchul Ahn (jack.ahn@hanwha.com), Intelligent C4I Team, Hanwha Systems Co.
- Received: 2023. 09. 26, Revised: 2023. 10. 12, Accepted: 2023. 10. 16.

## I. Introduction

지휘통제에서 지휘관은 임무 수행 달성을 위해 개인의 경험과 전문성에 의존한 의사결정을 하여 자신의 예하 부대에 임무를 하달한다. 지능형 지휘통제체계는 지휘관의 의사결정과 전장관리를 지원하는 핵심전력체계로 전장 인식, 지휘통제 등을 위해 정보처리 및 의사결정을 지원하는 자동화 체계이다[1]. 현재는 점점 복잡해지고 다양한 무기체계와 방대한 감시·정찰 자산의 데이터가 축적되어 지능형 지휘통제체계가 자동으로 중요 정보를 추출하고 융합하여 결과로 제공할 수 있도록 체계의 성능 개선이 필요하다. 인공지능(Artificial Intelligence, AI)에 기반한 지능형 지휘통제체계는 방대한 전장정보와 전술 데이터들이 학습 모델을 통해 분석되어 실시간으로 복잡한 전장상황 결과를 도출할 수 있다[1]. 이를 통해 지휘관은 군에 더 정확하고 신속한 임무 하달이 가능하고, 복잡한 미래 전장상황에 대비할 수 있다.

인공지능에 기반한 지능형 지휘통제체계를 구축하기 위해서는 우리 국군에 특화된 독자적인 인공지능 기술 및 태스크 수행이 필요하며 전장상황 분석용 인공지능 개발에 필요한 실 전장상황과 유사한 데이터셋 생성이 필요하다. 미국의 군 관련 연구개발을 담당하는 DARPA(Defense Advanced Research Projects Agency)는 COMPASS(Collection and Monitoring via Planning for Active Situational Scenarios)[2]나 AIDA(Active Interpretation of Disparate Alternatives)[3] 등의 여러 프로젝트의 개발을 진행하며 민간에서도 전장상황 판단 및 의사결정 지원을 위한 다양한 연구개발 프로젝트를 할 수 있도록 SM-KBP(Streaming Multimedia Knowledge Base Population) 등과 같은 많은 학회를 개최하여 오픈 데이터로 제공하고 있다. 국내에서는 군사작전 정보나 지휘통제 관련 데이터는 군에서 취급하는 기밀 자료로 열람할 수 있는 비취인가자가 아니면 전장상황을 설명할 수 있는 실 데이터를 발견하기 어렵다. 전장상황 분석용 인공지능을 개발하기 위해 활용할 수 있는 오픈 데이터는 웹에 공개된 뉴스 보도자료, 국방백서, 군사용어 사전 등의 데이터가 있다. 뉴스 보도자료는 훈련, 포격, 발사 등과 같은 특정 상황에 대해 자세하게 작성되어 있다. 전면전, 국지전 등과 같은 직접적인 전장 상황으로 활용하기에는 한계가 있지만, 군이나 민간에서 발생하는 전개 상황에 대한 정보로 활용할 수 있다. 뉴스 보도자료나 민간에서 취득할 수 있는 정보들은 오정보가 있을 가능성이 높기 때문에 이를 고려하여 데이터를 활용해야 한다.

본 논문에서는 지휘관에게 정확한 의사결정을 지원하기 위해 사전에 데이터셋 구축 방안을 제안한 선행연구 [4]의 다음 단계인 최종적인 결과를 도출하기 위한 데이터셋 생성 방안에 대해 설명한다. 기존에 연구한 선행연구 [4]는 전장상황 분석용 인공지능 개발에 필요한 전장상황 모의 시나리오 기반의 다양한 상황을 의미하는 가설 데이터셋을 생성하는 방안에 대해 제안하였다. 가설 데이터셋 생성은 4단계로 이뤄져 있으며 입력 부분부터 중간 단계인 2.5 단계까지 진행하였고, 데이터셋 생성을 위해 '전장상황용 학습 데이터셋 생성도구' SW를 개발하였다. 본 논문은 선행연구 [4]의 다음 단계부터 최종 출력 부분에 대해 설계된 내용과 개발한 부분을 기술한다. 최종적으로 지휘관에게 신뢰도가 높은 속성을 바탕으로 전장인식에 대한 결과를 제공하기 위해 인공지능 기반 상황분석 데이터셋 생성 방안에 대해 제안한다. 선행연구 [4]와 본 논문에서 제안한 방법을 기반으로 직접 구축한 데이터셋을 활용하여 여러 인공지능 모델들이 사전에 연구되었다. 여러 관련된 정보들이 계층적으로 연결된 각 가설 간의 관계를 반영한 계층적 그래프 신경망 기반 모델과 유사한 데이터를 서로 융합하는 군집화 모델, 그리고 최종적으로 전장상황 결과를 추천하는 모델을 설계 및 제안하였다[5, 6]. 인공지능 기반 지능형 지휘통제체계의 End-to-End 모델을 통해 결과를 출력하는 전체적인 데이터셋 구축 방안에 대해 제시하며 선행연구 [6]을 확장하여 데이터를 검증한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 연구와 연관된 용어와 관련 연구에 대해 기술하고 3장에서는 데이터셋 정의 및 설계 방안, 프로토타입 구현에 대해 기술한다. 4장에서는 구축한 데이터셋을 이용하여 선행연구 [6]의 추천 모델을 좀 더 확장하여 실험 설정 및 성능 검증, 결과를 제시하고, 5장에서 결론을 기술한다.

## II. Preliminaries

### 2.1 Related Works

#### 2.1.1 DARPA's AIDA Program

미국 DARPA의 AIDA 프로그램은 전장상황 판단 및 의사결정 지원을 위한 기술을 개발한 프로젝트이다. 실시간으로 전장상황 정보를 반영하여 신속하고 정확하게 지휘관의 의사결정을 지원한다. 전체 구조는 Fig. 1.과 같다[3]. AIDA 프로그램은 멀티모달 데이터(동영상, 텍스트, 음성 등)를 입력하여 인공지능을 통해 각종 상황 및 상황의 주변

정보들을 노드로 출력한다. Fig. 1.의 TA 3과 같이 출력값 노드를 다중가설(Multiple Hypothesis) 용어를 사용한다.

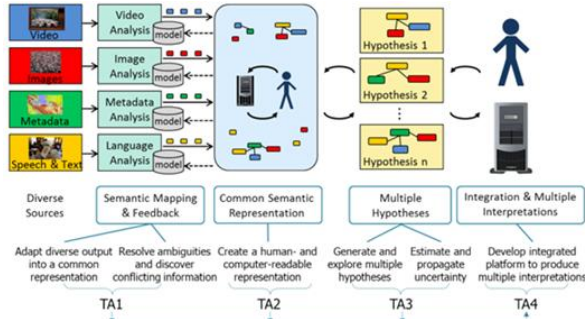


Fig. 1. DARPA's AIDA Program

2.1.2 Virtual Battlefield Situation Dataset

전장상황과 유사한 상황분석 데이터를 생성하기 위해 여러 군사 전문가와 의논하여 토픽 30개의 모의 시나리오 보고서를 구성하였다. 국내에서도 AIDA 프로그램의 다양한 상황 데이터인 다중가설과 같이 전장상황에 대한 여러 정보를 융합하고 분석하기 위해 가설 데이터셋 구축 방안을 제안하였다[4]. 가설 데이터 생성을 위해 전장상황에 대한 모의 시나리오를 설계 후 시나리오 문장의 사건/개체/관계 정보를 식별하여 사건과 사건 주변 정보를 노드로 표현하는 그래프 형태의 가설 데이터를 생성한다. 전장상황 모의 시나리오 기반의 다양한 사건을 단위가설(Unit Hypothesis), 후보가설(Candidate Hypothesis), 집합가설(Aggregation Hypothesis) 순으로 데이터셋 구축을 위한 학습 데이터셋 가설 생성도구 SW 설계 방안 및 구현 동작을 설명한다. 주요 용어 및 가설에 대한 정의는 Table 1.과 같다.

Table 1. Description of Terms and Datasets

Item	Contents
Event	Information that describes a specific action or affair occurring in the military
Object	Independent individual elements that describe the real world
Object-Object Relation	Independent individual elements that describe the real world
Event-Object Relation	Independent individual elements that describe the real world
Unit Hypothesis	A hypothesis consisting of one event and one or more objects, and an event-object relation
Candidate Hypothesis	A hypothesis that merged overlapping or similar unit hypotheses from multiple unit hypotheses
Aggregation Hypothesis	A hypothesis that separates conflicting candidate hypotheses and consists of a set of related candidate hypotheses

Table 1.의 단위가설은 사건(Event)과 하나 이상의 개체(Object) 그리고 사건과 개체를 연결하는 사건-개체 간 관계(Event-Object Relation)로 구성된다. 사건은 공격, 방어, 이동 등과 같은 정보이며 개체는 시간, 방향, 지역 등에 대한 정보이다. 후보가설은 단위가설보다 넓은 개념이며 여러 개의 유사한 단위가설이 합쳐진 가설이므로 후보가설을 먼저 생성 후 단위가설을 생성하는 방안에 대해 제시하였다. 또한, 후보가설 내용들은 서로 데이터가 충돌하지 않고 연관성 있는 가설로 군집화한 집합가설을 생성하는 방안에 대해 제안하였다[4].

2.1.3 Embedding Method of Hierarchical Data

모의 시나리오 보고서의 자연어 문장은 사건, 개체, 사건-개체 간 관계의 식별을 위해 자연어처리 모델을 거쳐 온톨로지 형태로 지식베이스에 구축된다. 쿼리를 통해 지식베이스에서 추출된 데이터는 자연어 단어 및 문장이기 때문에 전처리하는 과정이 필요하다.

온톨로지로 구성된 형태인 사건 및 개체별 유형은 Owl2Vec 모델을 활용하여 임베딩한다. Owl2Vec 임베딩 모델은 온톨로지의 클래스를 임베딩 하는 방법으로 계층적 구조에 대해 각 클래스의 유형을 구분하여 유형별 임베딩이 가능하다[7]. 따라서, 사건 및 개체별 유형은 단어 단위로 임베딩하여 사건에 대한 주변 개체 정보들을 보존할 수 있다. Stanford 대학에서 개발한 GloVe 임베딩 모델은 어휘벡터 모델로 함께 쓰이는 단어들에 대한 학습뿐만 아니라 전체 데이터셋의 통계 정보도 학습하여 연관된 단어 추정이 가능하다[8]. 지식베이스에서 추출된 데이터는 사건 및 개체별 유형 외에 각 사건 mention과 개체 mention이 존재한다. 각 Mention은 사건 및 개체 유형 데이터의 보고서나 문장의 원문, 관련된 정보들이 작성된 문장 또는 단어로 구성되므로 GloVe 모델을 활용하여 임베딩한다. GloVe 어휘 벡터 모델을 통해 mention과 연관된 단어들을 함께 임베딩 하여 각 단어 단위의 정도를 가깝게 생성할 수 있어 각 단어 간의 유사한 정확성이 높아 임베딩 방안으로 활용하였다[9]. GloVe 모델은 모의 시나리오 보고서 및 국방백서, 국방일보, 군사용어 사전, 국방과학기술용어, 뉴스 보도자료 등 말뭉치를 활용하여 사전 학습하였다.

Fig. 2.는 단위가설 임베딩 벡터 생성 방안이다. 각 Owl2Vec 모델과 GloVe 어휘 벡터 모델로 임베딩 전처리한 사건 및 개체별 유형의 임베딩 벡터와 사건 mention과 개체 mention의 임베딩 벡터는 그림과 같이 결합(Concatenate)한다.



Fig. 2. Embedding Method of Unit Hypo

Fig. 3.는 후보가설 임베딩 벡터 생성 방안으로 동일한 사건들로 이루어진 단위가설 1차원 임베딩 벡터의 평균을 계산하여 후보가설의 1차원 임베딩 벡터를 생성한다.



Fig. 3. Embedding Method of Candidate Hypo

## 2.2 End-to-End Model Architecture

지휘관의 요구나 질문에 관련된 가설 간의 관련성을 파악하고, 여러 지표를 바탕으로 최종적으로 신뢰도가 가장 높은 결과인 종합적인 점수와 순위를 제공하기 위한 데이터셋 구축 방안에 대해 설명한다. 지휘관이 작성한 요구나 질문은 질의(또는 사용자 질의)라고 한다. 질의에 대해 예시를 들자면, ‘금일 오전, A부대가 어디서 훈련하였는가?’라는 질의는 ‘{사건 : 훈련, 시간 : 금일 오전, 주체 : A부대, 장소 : ?}’의 후보가설로 변환할 수 있다. 시간, 주체, 장소 개체 유형은 사건에 대한 주변 노드로 그래프 형태의 후보가설을 설계할 수 있다. 질의 예시에서 장소에 대한 정보가 없으므로 이를 질의와 후보가설 데이터를 활용한 머신러닝으로 예측할 수 있다.

질의와 후보가설 입력을 기반으로 최종적인 점수 및 순위를 도출하는 과정인 학습 및 추론 단계 모델 구조는 Fig. 4.와 같이 구조화하고 학습모델 데이터셋 생성을 위한 세부적인 속성 및 항목들을 설계하였다. 학습 단계 첫 번째 과정의 입력 데이터는 후보가설 및 질의를 입력하여 첫 번째 모델인 회귀모델을 통해 후보가설과 질의에 대한 관련성 정도를 도출한다. 두 번째 과정은 회귀모델을 통해 도출된 후보가설과 질의의 관련성이 집합가설과 조합가설에 입력될 수 있도록 전처리 과정이 진행된다. 세 번째 과정은 Owl2Vec과 GloVe 모델을 이용하여 임베딩 벡터로 변환

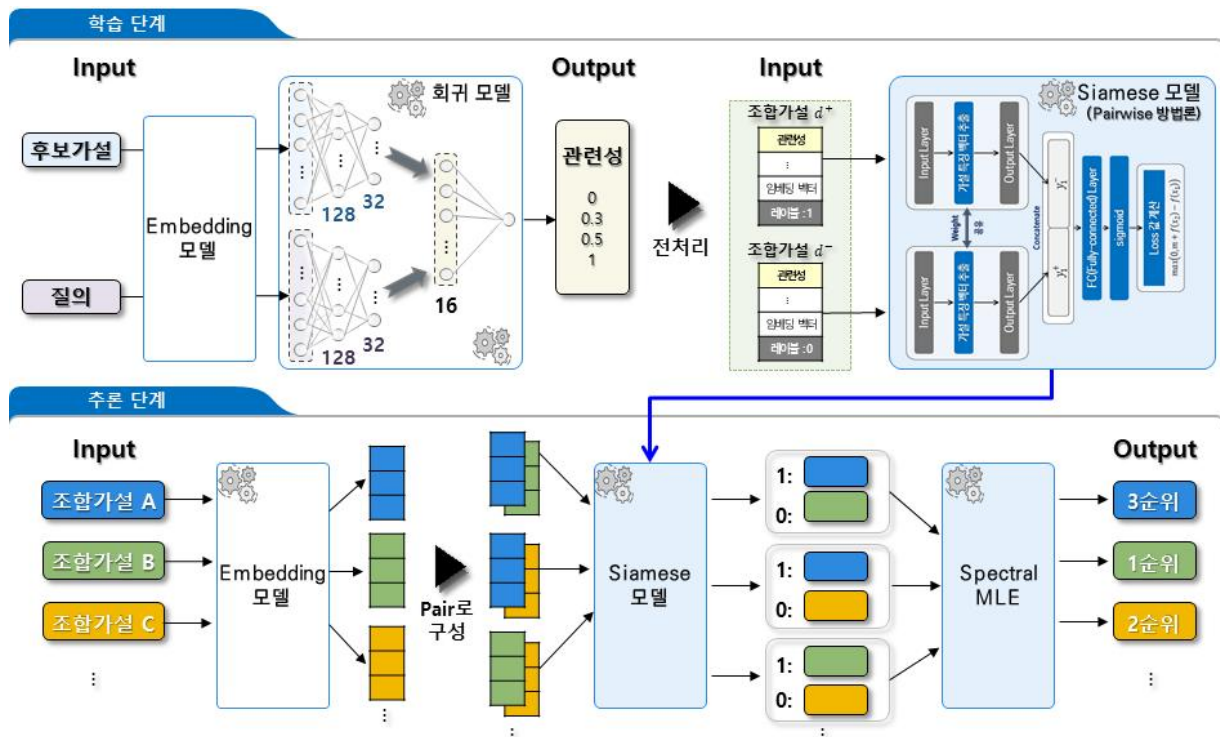


Fig. 4. Model Architecture

및 관련성 외 지표를 포함하고 있는 학습용 조합가설을 Pairwise 방법론인 Siamese 모델로 최적의 학습모델을 만든다. 추론 단계에서는 학습 단계와 같이 시험용 조합가설들을 임베딩 벡터로 변환하고 각 조합가설을 조합가설의 쌍으로 구성하여 Siamese 모델 및 Spectral MLE 방법을 이용하여 최종적인 종합순위 및 점수를 제공한다[6].

학습용/시험용 조합가설 데이터셋 구축 생성 방안은 다음 장에서 설명하며 본 절에서는 End-to-end Model 구조에 입력되는 데이터셋 구성 및 전처리, 학습모델에 관해 설명한다.

### 2.2.1 Attributes

최종적으로 지휘관에게 신뢰도가 가장 높은 결과 정보를 제공하기 위한 세부 속성 지표는 관련성(Relevance), 출처 신뢰도(Reliability), 정보 발생빈도(Information Occurrence Frequency), 지식 정확도(Knowledge Accuracy) 4가지로 각 가설 간의 중요 지표로 구성되어 있으며 각 속성에 대한 설명은 Table 2.와 같다.

관련성은 지휘관이 작성한 요구나 질문인 질의에 대해 후보가설과 관련있는 정도를 파악할 수 있는 지표로 최종적으로 종합점수 및 순위 결정을 위한 하나의 중요 지표다. 이를 예측하기 위해 사전에 지휘관들이 구축한 질의 데이터를 기반으로 후보가설의 데이터셋을 구축한다. Fig. 4.에서 질의와 후보가설을 임베딩 전처리한 후 회귀모델 또는 머신러닝 앙상블 모델을 이용하여 관련성 정도인 {0, 0.3, 0.5, 1} 출력값이 나오도록 분류 태스크를 수행한다. 질의 생성 방안과 머신러닝의 검증은 추후 연구를 통해 제시 및 검증한다. 출처 신뢰도는 전장자료 출처나 보고서에 대한 신뢰도 지표이며 정보 발생빈도는 후보가설에 포함된 단위가설들이 병합된 수를 계산하며 지식 정확도는 보고서에서 사건 및 개체를 추출하는 자연어처리 모델의 결과 정확도를 저장한 값이며 지식베이스에서 제공한다.

Table 2. Description of Attributes

Attributes	Description
Relevance	Score related to the query and candidate hypo
Reliability	Reliability score of battlefield data source
Information Occurrence Frequency	Based on aggregation hypo, # of candidate hypotheses merged (Likewise, the candidate hypo is the # of unit hypotheses merged)
Knowledge Accuracy	Accuracy of events and objects extracted from the knowledge base

### 2.2.2 Combination Hypothesis

조합가설은 분석 결과의 종합점수 및 순위를 산출하여 제공하기 위한 학습모델의 입력 데이터이다. Fig. 5.은 조합가설의 그래프 형태를 표현한 그림이다. 하나의 조합가설은 여러 개의 집합가설로 구성되어 있고, 계층적으로 집합가설과 후보가설도 동일하게 구성되어 있다. 단위가설은 하나의 사건과 여러 개의 개체로 이루어져 있다.

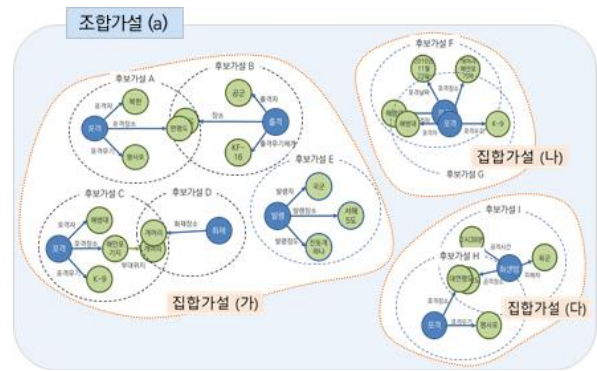


Fig. 5. Graph Structure of Combination Hypo

집합가설은 사용자가 군단이나 장소, 비슷한 시간대 등 연관성 있는 후보가설들을 2개 이상이 포함되어 만들어진 가설이며 집합가설에 있는 후보가설들은 서로 데이터 충돌 요소가 없는 가설(대립가설)이고 서로 연관 있는 가까운 후보가설로 묶여야 한다[4]. 또한, 조합가설에 있는 집합가설들도 서로 데이터 충돌요소가 없는 대립하지 않도록 조합하여 조합가설을 생성해야 한다. Fig. 5.의 조합가설 그래프 형태에 대한 정형 테이블인 조합가설의 하위 가설은 Table 3.과 같다. Table 3.에서 조합가설 (a)는 3개의 집합가설 (가), (나), (다)로 구성되어 있다. 집합가설 (가)는 5개의 후보가설 A, B, C, D, E로 구성되어 있다.

Table 3. Hierarchical Structure of Hypothesis

Hypothesis		
Combination Hypo	Aggregation Hypo	Candidate Hypo
(a)	(가)	A
		⋮
		E
	(나)	F
		G
	(다)	H
		I

Table 4.는 후보가설의 세부 속성이다. 후보가설 구조상 여러 개의 단위가설(사건+개체)로 구성되어 있지만 학습모델의 입력으로 처리하기 위해 앞서 설명한 ‘계층적 데이터셋 임베딩 방안’을 이용하여 임베딩 벡터를 도출한다. 따라서, 후보가설의 세부 속성은 임베딩 벡터값과 4가지 중요 지표(관련성, 출처 신뢰도, 정보 발생빈도, 지식 정확도)를 가진다. 집합가설의 세부 속성은 집합가설이 가지고 있는 4가지 지표와 여러 개의 후보가설, 그 후보가설을 이루는 임베딩 벡터값과 4가지 중요 지표를 가진다. 조합가설의 세부 속성은 여러 개의 집합가설과 그 집합가설을 이루는 임베딩 벡터값과 4가지 중요 지표로 구성되어 있다.

Table 4. Attributes of Candidate Hypo

Candidate Hypo	Attributes				
	Embedding Vector	Relevance	Reliability	Frequency	Accuracy
A					
B					
⋮					
I					

2.2.3 Pre-processing Method of Input Data

후보가설의 세부 속성들을 이용하여 집합가설을 생성하기 위해 집합가설에 포함된 후보가설 임베딩 벡터들과 4가지 지표의 평균을 취하고 집합가설의 4가지 지표를 결합(Concatenate)하여 하나의 집합가설 임베딩 벡터를 구성한다[6]. Fig. 6.은 집합가설 임베딩 벡터 생성 과정이다.

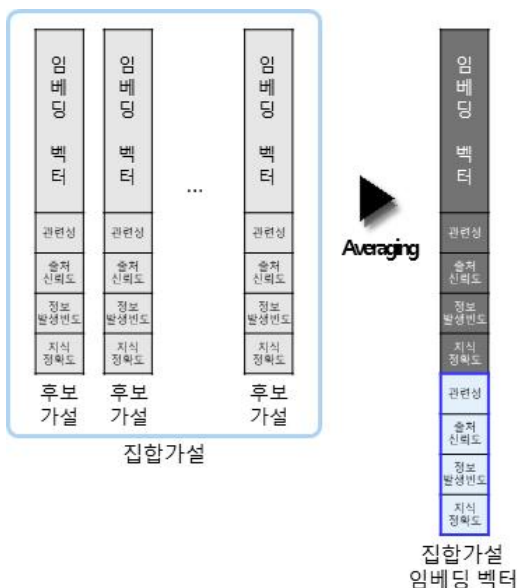


Fig. 6. Embedding Vector of Aggregation Hypo

Fig. 7.은 조합가설 임베딩 벡터 생성 과정이다. 집합가설의 임베딩 전처리 과정과 동일하게 각 조합가설 임베딩 전처리 과정에도 적용하여 조합가설에 대한 임베딩 벡터를 생성하여 학습모델에 활용한다. 조합가설에 포함된 집합가설 임베딩 벡터들의 평균을 취하고 조합가설의 4가지 지표를 더하여 하나의 조합가설 임베딩 벡터를 구성한다. 최종적으로 조합가설 임베딩 벡터가 모델의 입력값이 된다.

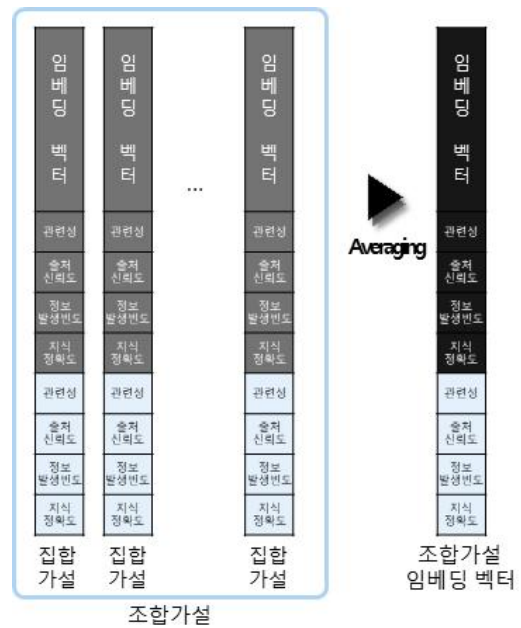


Fig. 7. Embedding Vector of Combination Hypo

2.2.4 Recommend Model Approach

추천 모델은 Loss 함수에 따라 Pointwise, Pairwise, Listwise 방법론으로 구분된다. Pointwise 방법론은 학습시 각각 독립적으로 학습 후, 추천시 점수 출력 후 점수별로 정렬하여 순위를 도출하며 Pairwise 방법론은 두 개의 조합가설을 한 쌍(Pair)으로 하여 우선순위를 고려하고, 각 가설 중 어떤 가설이 우선인지 각 속성 및 특성을 반영하여 학습할 수 있는 방법론이다. Listwise는 두 개의 조합가설을 한 쌍의 Pair로 하여 우선순위를 고려하고 특성을 반영하여 학습하는 방법론이다. Pointwise 방법론은 하나씩 독립적으로 학습하여 다른 순위를 고려하는 특성을 반영하지 못하고, Listwise 방법론은 N개의 조합가설 데이터가 있으면 1부터 N까지 각 순위를 레이블 작업하여 학습 데이터셋을 구축하기에 한계가 있다. Pairwise 방법론은 조합가설 쌍 중 어느 가설이 우선인지 중요 지표들을 보고 레이블(조합가설a > 조합가설b) 데이터셋으로 구축하기 쉽고, 순위의 특성을 반영하여 학습하기 때문에 Pairwise 방법론의 학습모델을 선택하여 설계한다[10].

### 2.2.5 Siamese Network Model

Siamese 모델은 두 개의 입력에 대해 독립적으로 신경망을 통과시키고, 2개의 신경망에 동일한 가중치(Weight)를 공유하여 두 개의 출력을 비교하는 구조의 Pairwise 방법론 추천 모델이다[11]. Siamese 모델을 선택한 이유는 두 개의 입력값이 독립적으로 학습되지만 각 가설 중 어느 가설이 더 정답에 가까운 정도의 수치를 Loss 함수를 이용하여 거리(Distance) 계산으로 우선순위 판별이 가능하므로 모델로 선정하였다. 또한, 쌍(Pair) 형태의 데이터에 대해 특정 모델에 대한 가정 없이 활용할 수 있는 학습모델이며 2개 중 우선순위를 출력하는 알고리즘으로 조합가설 데이터를 이용한다. 쌍 형태로 구축하고 학습모델에 입력하여 성능을 측정하기에 적합하다. 추론 과정에서 시험용 조합가설 데이터셋에 대해 쌍 형태의 조합가설 2개 중 어떤 가설이 우선인지 단순 두 개의 입력값을 비교하는 것이 아닌 후처리 과정인 Spectral MLE 알고리즘으로 전체 시험용 조합가설에 전체 순위를 도출할 수 있다[12].

## III. The Proposed Scheme

본 논문에서는 DARPA의 AIDA 프로그램 가설의 용어 및 용어의 의미는 동일하게 활용하지만 데이터셋 구성은 다르게 구성하였다. Table 5.는 가설별 설명이며 가설 데이터셋은 총 4가지로 구성된다[9].

Table 5. Description of Hypotheses

Hypothesis	Description
Unit Hypothesis	A hypothesis consisting of one event and one or more objects, and an event-object relation
Candidate Hypothesis	A hypothesis that merged overlapping or similar unit hypotheses from multiple unit hypotheses
Aggregation Hypothesis	A hypothesis that consists of a set of related candidate hypotheses
Conflict Hypothesis	A hypothesis of a relationship in which data conflict elements exist between aggregation hypothesis
Combination Hypothesis	A hypothesis that separates conflicting Aggregation hypotheses and explains a comprehensive battlefield situation by combining aggregation hypotheses

가장 기본이 되는 가설인 단위가설은 한 사건에 대해 관련된 개체들과 사건과 개체 사이에 관계로 이루어진 가설

을 의미한다. 단위가설은 각 단계를 거쳐 후보가설, 집합가설 그리고 조합가설로 구성한다. 본 장에서는 선행연구 [4]의 단위가설, 후보가설, 집합가설 생성 방안 이후 집합가설과 대립가설을 활용한 조합가설 생성 및 조합가설 쌍 생성 설계 방안을 기술하고 최종적으로 지휘관 및 사용자에게 전장인식 결과를 제공하는 모델에 적용하기 위한 상황분석 데이터셋 생성 프로토타입 설계 및 SW 구현에 관해 설명한다.

### 3.1 Design Method for Data Construction

조합가설을 구축하기 위해서는 집합가설별 대립가설 정보가 필요하다. 대립가설 정보는 지식베이스에 구축한 개체 인스턴스별 OWL, SWRL 등 규칙 정의나 추론 엔진을 통해 대립 정보를 가져오거나 사람이 직접 대립 정보를 생성할 수 있다. 조합가설을 생성하기 위해서는 집합가설별 대립 정보를 배제하고 생성하는 것이 필수조건이며 집합가설 2개 이상을 포함하도록 조합(Combination)의 수만큼 조합하여 조합가설을 생성한다.

수식 (1)은 조합가설을 생성할 수 있는 조합가설의 최대 생성 개수이다. 여기서 최대 개수란, 대립 정보가 없을 경우의 최대 생성할 수 있는 개수를 의미한다. 대립 정보가 있으면 수식 (1)의 계산한 수보다 적다. 집합가설의 수가 4개이고 서로 대립인 집합가설이 없다고 가정했을 때 조합가설을 구성하는 집합가설 4개의 조합의 수는 최대 11개( ${}_4C_2 + {}_4C_3 + {}_4C_4 = 6 + 4 + 1 = 11$ )가 생성이 되며 하나의 집합가설이 대립가설이 하나라도 있는 경우에는 조합가설이 11개보다 적게 생성된다.

$${}_n C_2 + {}_n C_3 + \dots + {}_n C_n \quad (1)$$

Fig. 8.은 집합가설 간 대립 정보를 이용한 조합가설 생성 및 학습모델의 입력 데이터인 조합가설 쌍 생성 과정을 나타낸다. Fig. 8.에서는 집합가설이 총 4개인 (가), (나), (다), (라)로 구성되어 있다. 집합가설 (가)의 대립가설은 집합가설 (나)이므로 두 개의 집합가설 (가), 집합가설 (나)는 같은 조합가설로 생성하지 않는다. Table 6.은 집합가설 4개에 대해 조합가설의 최대 생성 개수 11개의 목록이며 집합가설 (가)의 대립 정보를 이용하여 생성 여부를 확인할 수 있다. Table 6.에서 집합가설 (가)와 집합가설 (나)는 서로 대립이므로 생성되지 않는다. 생성되지 않는 개수 4개를 제외한 7개의 조합가설을 생성할 수 있다. 또한, Fig. 8.에서 집합가설 (나)의 대립가설은 집합가설 (가)와 집합가설 (다)이므로 집합가설 (나)는 집합가설 (가)와

집합가설 (다)가 포함되지 않도록 조합가설을 생성해야 한다. Table 6. 이후 조합가설이 생성될 수 있는 7개 중 집합가설 (나)와 (가) 또는 (다)가 포함되지 않도록 한다. Table 7.은 집합가설 (나)의 대립 정보를 이용한 생성 여부 목록이다.

Table 6. Combination for Aggregation Hypo (가)

Index	Combination for Aggregation Hypo (가)	Generate
#1	(가), (나)	X
#2	(가), (다)	O
#3	(가), (라)	O
#4	(가), (나), (다)	X
#5	(가), (나), (라)	X
#6	(가), (다), (라)	O
#7	(가), (나), (다), (라)	X
#8	(나), (다)	O
#9	(나), (라)	O
#10	(나), (다), (라)	O
#11	(다), (라)	O

Table 7. Combination for Aggregation Hypo (나)

Index	Combination for Aggregation Hypo (나)	Generate
#1	(가), (다)	O
#2	(가), (라)	O
#3	(가), (다), (라)	O
#4	(나), (다)	X
#5	(나), (라)	O
#6	(나), (다), (라)	X
#7	(다), (라)	O

집합가설 (다)와 (라)에 대해서도 대립 정보를 이용하여 조합 생성 여부를 판단하며 Table 8.과 같이 5개의 조합가설이 생성된다. 최종적으로 11개의 집합가설 중 집합가설별 대립 정보를 배제하여 조합한 가설이 생성되는 조합가설 수는 총 5개이다.

Table 8. Construction of Combination Hypo

Combination Hypo	Aggregation Hypo
a	(가), (다)
b	(가), (라)
c	(가), (다), (라)
d	(나), (라)
e	(다), (라)

조합가설 생성 이후, Pairwise 방법론의 Siamese 모델을 활용하기 위해 앞서 생성된 조합가설들을 가져와서 조합가설 쌍(Pair)을 생성해야 한다. 전체 조합가설의 개수가  $n$ 개라고 가정한다면, 쌍별로 2개를 입력 데이터로 구성해야 하므로 전체  $n$ 개 중 2개를 조합( ${}_n C_2$ )하여 조합가설 쌍을 생성한다. Table 8.의 5개의 조합가설에 대해 조합가설 쌍을 계산하면 Fig. 8.에서 조합가설 쌍 생성 단계 이후 총 10개( ${}_5 C_2 = \frac{5 \times 4}{2 \times 1} = 10$ )를 생성할 수 있다. Fig. 8.에서 조합가설 쌍을 표기하는  $d^+$  와  $d^-$  의 의미는 조합가설  $d^+$  는 조합가설  $d^-$  보다 우선(선호)한다는 것을 의미하며 Index #1에서 조합가설 a는 조합가설 b보다 우선하는 것을 의미한다. 최종적으로 학습모델에 입력되는 데이터셋은 임베딩 벡터, 집합가설 및 후보가설의 각 4가지 지표를 포함한 조합가설 2개로 묶인 조합가설 한 쌍이다.



Fig. 8. Construction Process for Pair of Combination Hypo using Aggregation Hypo



### 3.2 Building a Prototype for Data Construction

대립적인 요소가 있는 집합가설들은 배제하며 조합하여 도출한 결과를 생성하기 위한 전장상황분석 학습 데이터셋 생성도구 SW의 프로토타입 구조는 Fig. 9.와 같이 구현하여 학습 데이터셋 구축을 진행한다. 데이터 생성 구축 방안은 상위 태스크인 2가지 부분으로 구분된다. 집합가설의 대립 정보 관리 및 조합가설 생성 부분과 조합가설 쌍 생성 부분으로 기능이 나뉜다. 각 기능은 세부 기능으로 나뉘서 데이터셋이 구축된다. 전장상황분석 학습 데이터셋 생성도구 SW 개발 환경은 JAVA 및 python 언어를 기반으로 웹 서버 및 UI를 구현하였고, MySQL을 이용한 DB를 구축하여 학습 데이터셋 생성을 위한 시스템 환경을 구성하였다.

#### 3.2.1 Construction for Combination Hypothesis

학습 데이터셋 생성도구 SW 시스템은 '집합가설 별 대립 정보 생성 및 관리' - '조합가설 속성 생성' - '조합가설 데이터 관리' 순으로 진행된다.

'집합가설 별 대립 정보 생성 및 관리' 과정에서 대립 정보는 기존 선행연구 [4]에서 진행한 유사/식별 모델의 대립 정보를 가져와서 집합가설을 생성하거나 외부 속성 정보 DB를 통해 대립하는 집합가설의 정보를 가져온다. 외부 속성정보 DB에서 지식베이스 내 OWL, SWRL 등 규칙 정의 또는 추론 엔진을 통해 추론된 결과인 대립가설 정보가 제공된다. 집합가설 및 대립가설 정보를 가져온 이후, 집합가설 군집화를 진행하고 집합가설 군집화는 조합가설을 생성하는 과정을 의미한다.

먼저, 집합가설의 대립 정보를 배제하는 것이 필수조건이므로 이를 만족시키며 집합가설 2개 이상을 포함하도록 조합의 수만큼 조합하여 조합가설을 생성한다. Fig. 10.에서 우측 상단 '조합가설 생성' 버튼을 클릭하면 '집합가설 조합 개수'를 입력할 수 있는 윈도우를 보여준다. 집합가설 조합의 개수를 선정하여 선정한 조합의 수만큼 조합하거나 그 이하의 조합의 수를 이용하여 조합가설을 생성할 수 있다. 집합가설 조합의 수를 5개 선택했다면 5개 씩 조합하도록 묶거나 2개 이상, 5개 이하의 조합의 수만큼 집합가설을 군집화한 조합가설 데이터셋이 생성된다. 집합가설 군집화된 결과인 조합가설을 확인할 수 있다. 집합가설 대립 정보를 가져오고 조합의 개수를 직접 선정 후 시스템에서 자동으로 계산되어 있어 자동화된 조합가설 생성 레이블 과정이 진행된다. Fig. 11.은 집합가설 군집화 결과인 조합가설 생성된 결과 화면이다.

'조합가설 속성 생성' 과정은 속성 지표 중 출처 신뢰도, 정보 발생빈도, 지식 정확도 등은 자동으로 반환된 값을 가져오고 관련성 데이터는 사람이 직접 수치를 입력하여 레이블 작업을 하거나 Fig. 4.의 입력 부분의 후보가설 및 질의의 회귀모델을 통해 예측하여 전처리 후 조합가설의 관련성 값을 대체할 수 있다. 시스템 상 회귀모델을 통해 추론되는 값으로 대체하는 기능이 없으므로 Fig. 9.와 같이 '속성 정보 가져오기'와 같이 관련성 값 정보를 가져오거나 '속성 환경설정'을 통해 랜덤값 지정 또는 '속성값 직접 입력' 기능을 통해 속성값 레이블을 직접 입력하고 저장할 수 있다. Fig. 12.는 관련성 속성 정보를 직접 입력

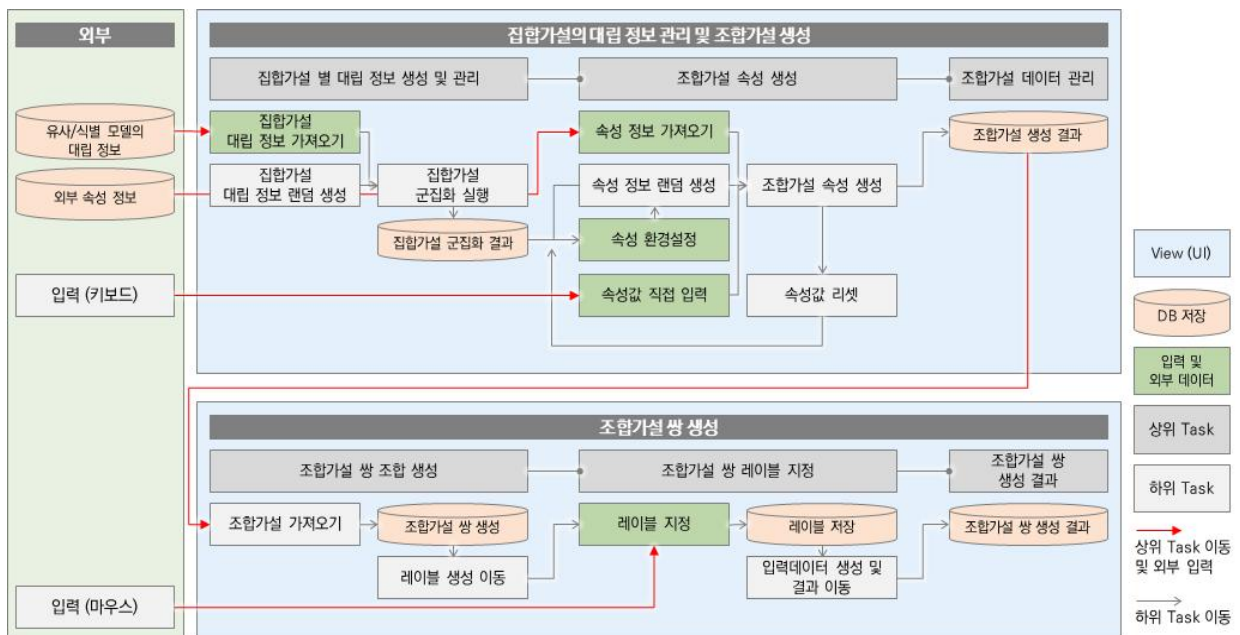


Fig. 9. Building a Prototype to Construct the Dataset for Battlefield Situation Analysis

및 저장, 삭제할 수 있는 기능과 조합가설 쌍이 생성된 화면이다. ‘조합가설 쌍 생성’ 시스템 흐름은 다음 절에서 설명하며 관련성 속성 정보 입력은 조합가설 쌍 별 집합가설의 내용을 직접 확인하여  $d^+$ 에  $d^-$ 보다 높은 관련성 값을 입력 후 저장할 수 있으며 조합가설 쌍의 행을 삭제하는 기능을 넣어 ‘조합가설 데이터 관리’가 가능하다.

### 3.2.2 Construction for Pair of Combination Hypo

학습 데이터셋 생성도구 SW 시스템은 ‘조합가설 쌍 조합 생성’ - ‘조합가설 쌍 레이블 지정’ - ‘조합가설 쌍 생성 결과’ 순으로 진행된다.

학습모델은 조합가설 두 개를 비교했을 때 어떤 것이 우선인지 판단하는 모델이기 때문에 입력 데이터가 조합가설 2개 중 어느 가설이 우선순위( $d^+ > d^-$ )인지 형태의 데이터셋을 생성해야 한다. ‘조합가설 쌍 조합 생성’ 과정은 전체  $n$ 개 중 2개를 조합( ${}_n C_2$ )하여 조합가설 쌍을 생성하는 자동화 과정을 거쳐 결과를 반환한다. Fig. 11.에서 우측 상단 ‘조합가설 레이블 지정’ 버튼을 클릭하면 Fig. 12.와 같이 앞서 설명한 ‘조합가설 속성 생성’의 관련성 정보 입력 및 전체 개수 중 2개를 조합한 조합가설 쌍이 생성된 결과인 ‘조합가설 레이블 지정’ 화면을 볼 수 있다. ‘조합가설 쌍 레이블 지정’ 과정은 사용자가 직접 조합가설 2개로 묶인 한 쌍 중 1개의 우선순위라고 판단되는 것을 선택하여 직접 레이블 작업 후 저장하거나 Table 9.과 같은 알고리즘으로 형식으로 자동화하여 조합가설 쌍의 레이블이 진행된다.

레이블 자동화 알고리즘에서 출처 신뢰도는 관련성 다음으로 중요 지표이다. 출처 보고서의 신뢰도에 따라 오류일 가능성을 가려낼 수 있기 때문이다. 예를 들어, 인간보고서라는 사람이 작성한 출처에서 사람이 눈으로 확인했을 때 비행기인지 새인지 구별하기 어렵고, 영상보고서가 출처일 경우에는 직관적으로 어떤 물체인지 비교적 정확하게 확인할 수 있어 신뢰도 자체가 높다. 출처 신뢰도는 사람이 직접적으로 판단하는 경우에 출처 신뢰도 자체에 오류가 있을 가능성이 크고 영상 정보나 각종 근거 자료를 보고 분석하는 것이 정확하므로 출처 신뢰도가 0.3 이상 차이 날 경우, 출처 신뢰도가 높은 조합가설이 우선순위 조합가설이라고 볼 수 있다. 정보 발생빈도가 후보가설 기준, 유사한 단위가설들이 병합한 수치이다. 여러 보고서에 유사한 내용이 많이 나올수록 단위가설 및 후보가설이 병합할 가능성이 높아 정보 발생빈도가 높아지므로 해당 보고가 오류일 확률은 비교적 낮아진다. 지식 정확도는 사건 및 개체를 추출하는 모델의 결과를 지식베이스에서 함께

저장하여 그 반환된 결과를 가져와서 활용하지만, 지휘관의 입장에서 모델의 정확도보다 실제 작성된 보고서 기반으로 출처 신뢰도나 생성된 보고서 단어의 정보 발생빈도를 좀 더 활용할 거라 판단하여 지표 중 가장 후순위로 두고 자동화 조합가설 쌍 생성 레이블을 구현하였다.

Table 9. Pseudo Code for Auto-Label Algorithm

if	When the reliability differs by more than 0.3 :
	Hypo with high reliability take priority
else	
if	When the frequency is the same :
	Hypo with the largest # of candidate hypo take priority
else if	When the # of candidate hypo is the same :
if	When the knowledge accuracy is high :
	Hypo with high knowledge accuracy take priority
else	
	Commanders Judgment
else	
	Hypo with a high frequency or a large # of candidate hypo take priority

출처 신뢰도가 0.3 이상 차이, 정보 발생빈도나 후보가설 개수 차이가 0(same) 등 임계치 수치는 시스템의 환경 설정에서 임계치 수치를 변경할 수 있어 사용자에게 맞는 자동화 레이블이 가능한 조합가설의 쌍 데이터셋을 생성할 수 있다. Fig. 13.은 조합가설 쌍 생성을 위해 Fig. 12.와 같이 직접 조합가설 쌍 하나씩 우선인 것을 사람이 직접 레이블 작업을 하거나 Table 9.의 알고리즘으로 자동으로 레이블을 수행한 결과인 ‘조합가설 쌍 생성 결과’ 화면이다. Fig. 14.는 학습모델의 입력값으로 수정이 필요하거나 삭제할 수 있는 기능을 추가하여 데이터셋 관리가 가능한 화면이다. Fig. 15.는 레이블을 수행한 결과인 조합가설 쌍 데이터셋을 다양한 형식으로 저장이 가능한 화면이다. 다양한 .json, .txt, .rdf 확장자로 저장이 가능하여 학습모델의 입력 데이터로 활용한다.

## IV. Evaluation

### 4.1 Datasets

지휘관이 전장상황을 파악하기 위해서는 전장 지식요소 중에서 연관된 정보만 선택하여 분석할 수 있어야 한다[4]. 따라서, 관련된 정보를 선정하기 위해 군사 전문가가 직접

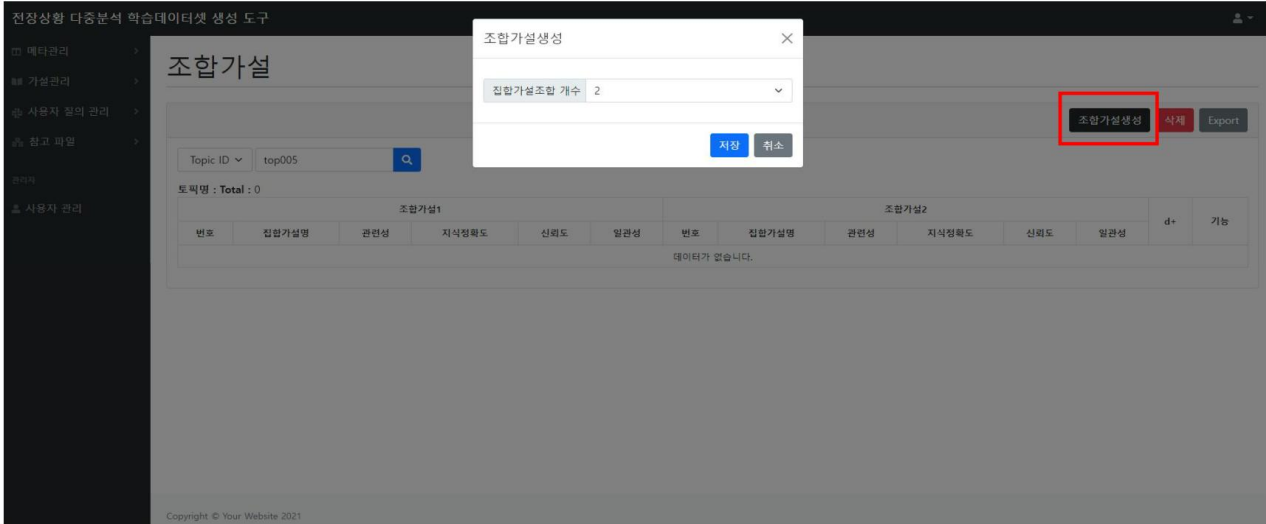


Fig. 10. Dataset Generator SW UI – Setting the Number of Combinations to Construct the Combination Hypotheses

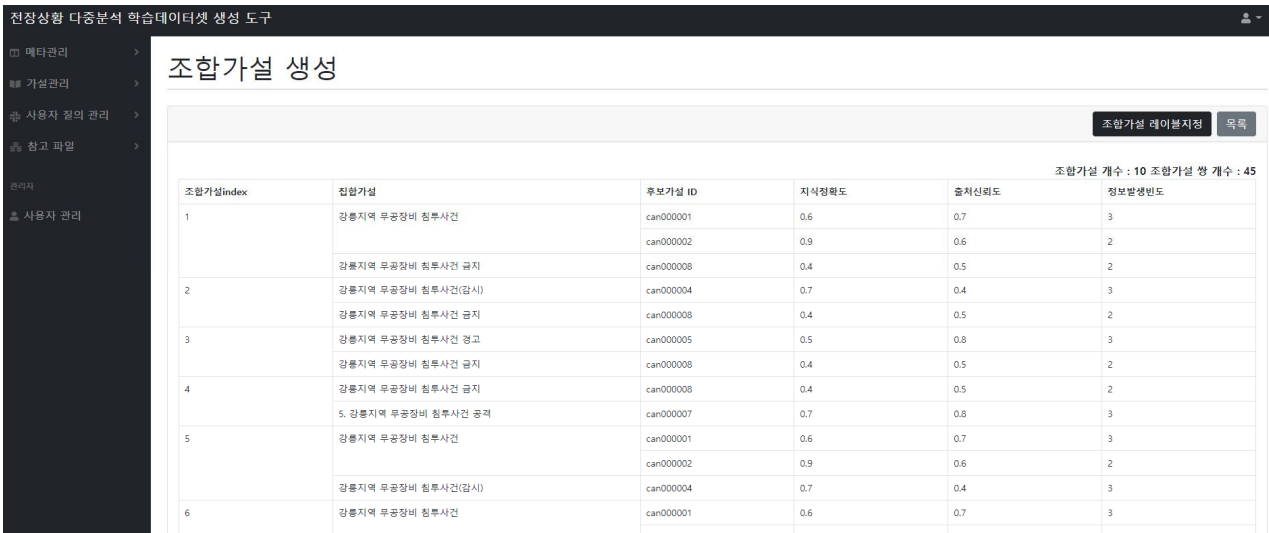


Fig. 11. Dataset Generator SW UI – Result of Construction the Combination Hypotheses



Fig. 12. Dataset Generator SW UI – Functions of Labeling the Combination Hypothesis and Relevance Attribute

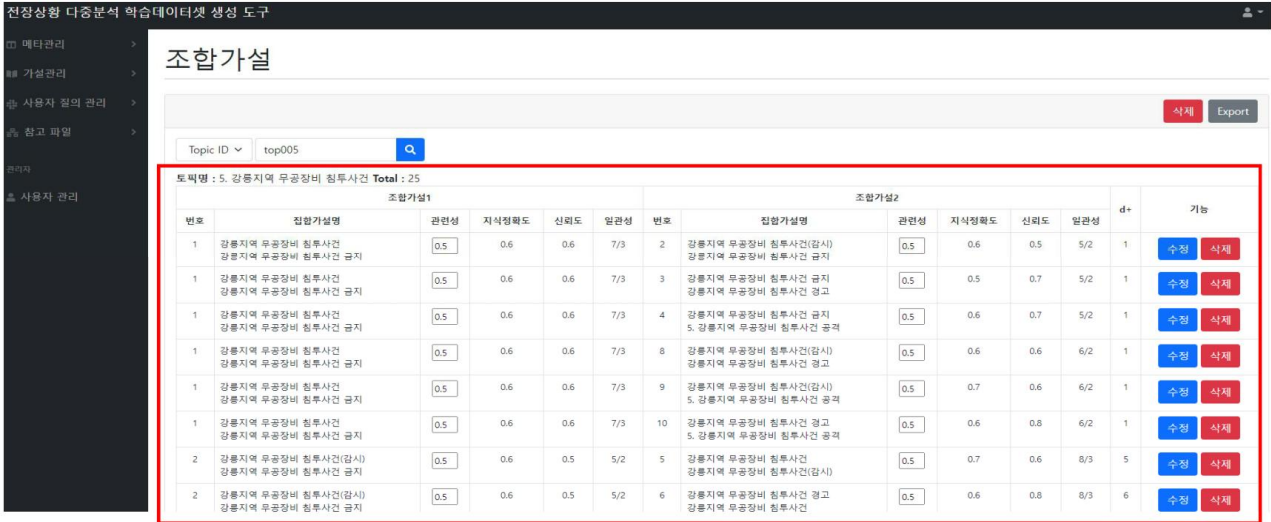


Fig. 13. Dataset Generator SW UI – Result of Construction the Pair of combination hypotheses

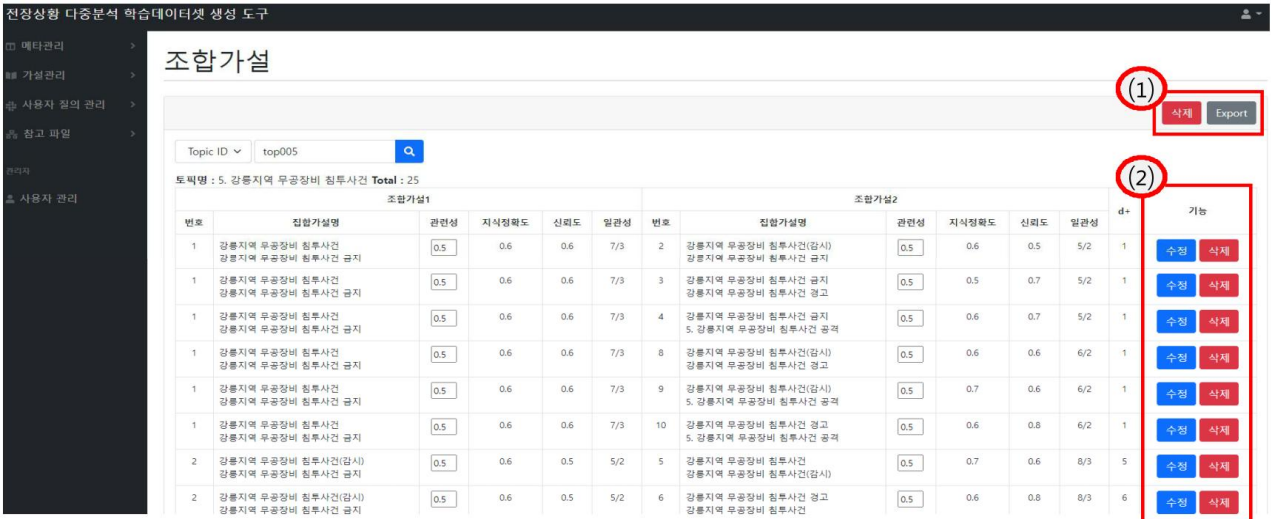


Fig. 14. Dataset Generator SW UI – Edit and Delete Functions

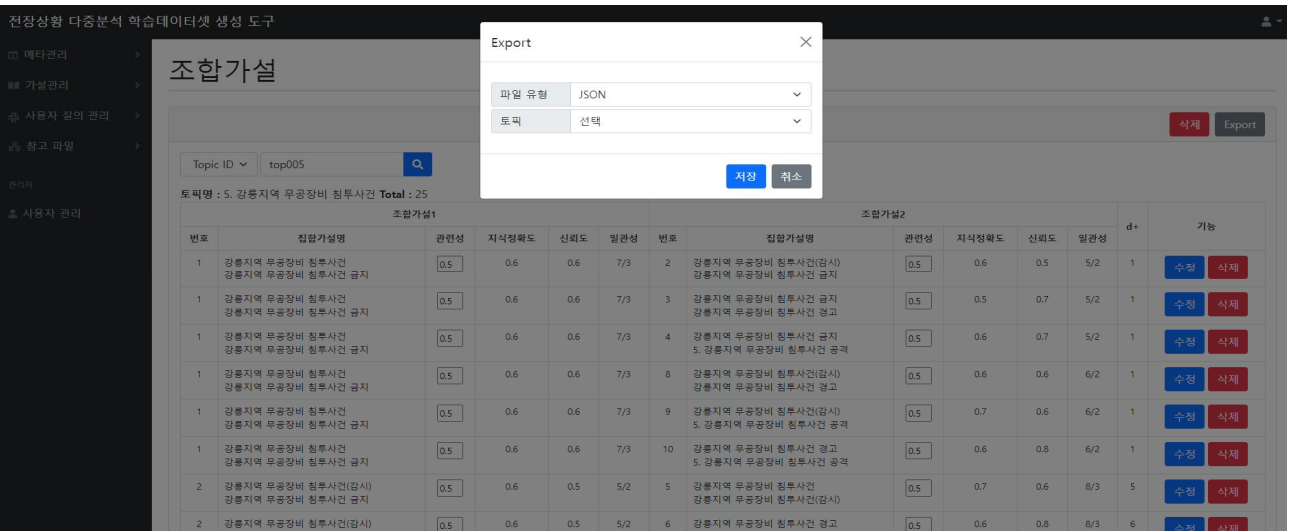


Fig. 15. Dataset Generator SW UI – Save and Export Functions

수동으로 9개의 토픽을 선정하여 집합가설 레이블 작업을 진행하였고 그에 맞는 조합가설, 조합가설 쌍 생성을 시스템에서 자동화하여 생성하였다.

Table 10.은 토픽 별 각 조합가설 쌍 개수를 나타낸다. 각 토픽의 데이터셋은 학습용 데이터셋과 시험용 데이터셋에 대해 80%, 20%를 적용하여 진행하였다. 시험용 데이터셋에 대해 군사 전문가가 직접 우선순위를 선정하였고, 학습모델을 통해 순위를 검증하였다. 학습모델의 입력 데이터는 조합가설 데이터셋으로 조합가설별 주요 지표, 순위 정보를 포함하고 있으며 학습모델에는 조합가설 쌍을 입력한다.

Table 10. Number of Datasets

Topic	# of Pair of Combination Hypo
#1	962
#2	176
#3	6790
#4	6634
#5	11274
#6	6562
#7	6902
#8	4428
#9	6870
Total	50,598

#### 4.2 Ranking Model and Post-process Method

Fig. 4.와 같이 학습 단계에서 Pairwise 방법론의 Siamese 모델을 이용한 학습모델을 생성하고, 추론 단계에서 Siamese 모델의 쌍(Pair) 형태의 출력값을 Spectral MLE를 활용하여 최종적인 순위 결과를 도출한다.

Spectral MLE 는 총 3단계로 진행한다[12]. 1단계는 Pairwise 순위의 행렬 변환 과정이고, 2단계는 Random Centrality를 활용한 초기화 단계를 수행한다. 3단계는 초기화 단계 이후 값들을 근사하는 정제(Refinement) 과정을 진행하여 최종값을 도출한다. Fig. 16.의 예시와 같이 조합가설 A, B, C, D에 대해 학습모델을 통해 출력된 우선순위 쌍 형태로 가정하고, 예시를 통해 Spectral MLE의 후처리 방법을 단계별로 설명한다.



Fig. 16. Example of Siamese Model Output

조합가설 A와 D, 조합가설 B와 D 중 모델의 결과로 어느 가설이 우선인지 판단할 수 없지만 후처리 알고리즘인 Spectral MLE을 통해 전체 순위를 추론할 수 있다. 1단계는 Pairwise 순위로 출력된 출력값들을 전체 Listwise 순위로 변환하기 위한 초기 단계로 변환하는 과정이다. 먼저, 조합가설의 개수( $n$ )에 따라  $n \times n$  형태의 우선순위 행렬  $M$ 을 생성하고 조합가설에 해당하는 행에 따라 해당 행의 값에 우선이면 1 값을 부여하고, 아니면 0 값을 부여한다. 각 행에 값이 채워진 후 초기화 과정을 진행하는데 수식 (2)의 전이행렬  $P$ 를 구성한다[6].

$$P_{ij} = \begin{cases} \frac{1}{d_{\max}} M_{ij} & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{k \neq i} M_{ik} & \text{if } i = j. \end{cases} \quad (2)$$

Fig. 17.과 Fig. 18.은 Pairwise 형태의 행렬 변환 및 초기화 과정을 나타낸다. Fig. 17.의 ① 은 조합가설 4개에 따라 조합가설 A가 조합가설 B와 C보다 우선이 아니므로 (1,2)과 (1,3)에 0 값을 부여하고, (2,1)과 (3,1)에 1 값을 부여한 결과이다. 마찬가지로 모든 행렬 성분에 대해 우선이면 1, 아니면 0을 부여한다. Fig. 17.의 ② 는 초기화 과정을 진행한 전이행렬  $P$ 의 결과이다. 초기화 수행 후  $P$  행렬은 각 열의 총합이 1이 되도록 대각노드를 계산하고 대각노드를 제외한 노드는 0으로 채운다. Fig. 18.은 대각노드 계산 및 남은 노드를 0으로 대체한 결과 행렬을 나타낸다.

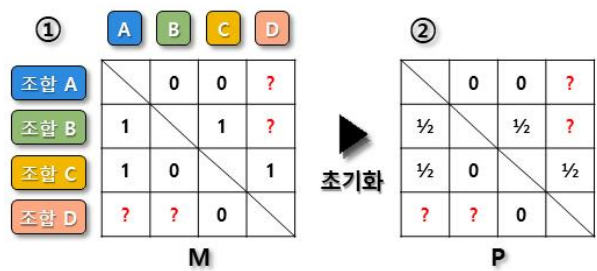


Fig. 17. [Step 1-1] Example of Pairwise Output to Matrix Conversion and Initialization

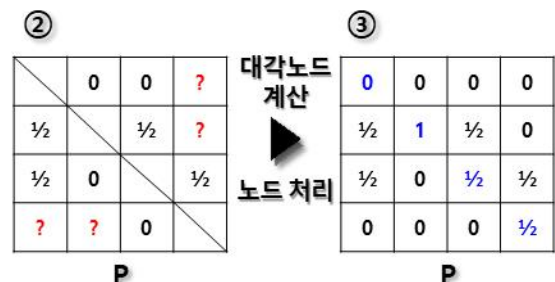


Fig. 18. [Step 1-2] Example of Matrixization

2단계는 값들을 구하기 위한 Random Walk 형태로 변환 및 고정분포를 계산하는 Random Centrality 단계이다. Random Walk란 그래프 노드(하나의 조합가설)를 무작위로 이동하는 경로를 의미하고 마코브 체인의 상태변이 확률로 계산한다. 고정분포란 각각의 노드가 가장 우선되는 확률값을 의미한다. Centrality(중심성)은 일반적으로 네트워크 그래프에서 노드의 중심성을 파악하고 계산하는 방법으로 Spectral MLE를 수행하기 위해 Random Centrality를 수행한다[13]. 쌍별 관찰을 기반으로 마코브 체인의 Random Walk를 반복적으로 수행한다. Fig. 19.는 P 행렬을 그래프로 변환한 Random Walk를 나타낸다.

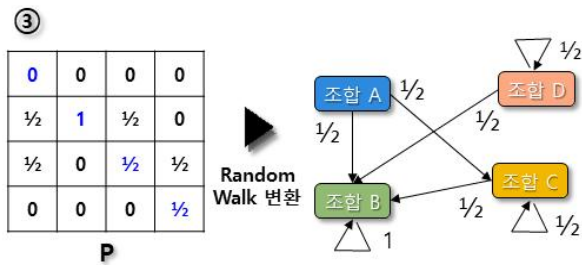


Fig. 19. [Step 2-1] Example of Random Walk

이후에 전이행렬 P의  $P^T$ 를 구성하여 고유벡터  $\pi$ 를 계산한 고정분포를 반환한다. Random Walk에 자주 머물거나 고정분포의 값이 높을수록 선호도가 높다고 할 수 있다. 고정분포  $\pi$ 는 3단계 정제 과정의 초기값(seed)으로 사용한다. Fig. 20.은 Random Walk를 반복적으로 수행한 이후 노드가 우선되는 확률 고유벡터  $\pi$ 인 고정분포를 나타낸다.

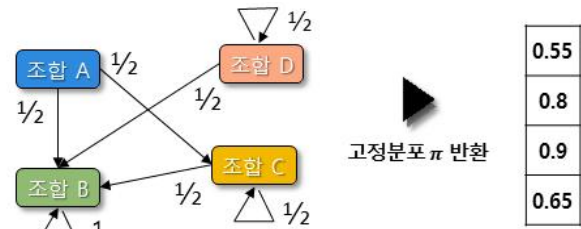


Fig. 20. [Step 2-2] Example of Stationary Distribution

3단계는 고정분포  $\pi$ 의 근사적 초기값으로 활용한 출력 값들의 정제하여 순위를 도출하는 단계이다. Fig. 21.는 정제 과정을 나타낸 예시이다. 2단계의 고정분포  $\pi = [0.55, 0.8, 0.9, 0.65]$ 는 각 조합가설의 Score(점수) 값이며 초기값으로 활용한다. 시행 1의 초기값을 이용하여 선호도(우선)를 확인하면 기존 모델의 출력값인 (B 0.8 > A 0.55), (C 0.9 > A 0.55) 선호도와 Score 값이 위배되지 않으므로 시행 1의 조합가설 A Score 값인 0.55가 변경되지 않는다. 그다음 시행은 이전 시행에서 업데이트가 되었으면 업데이트가 된 Score 배열을 가져오거나 업데이트가 되지 않았으면 이전 시행의 Score 배열을 활용한다. 시행 2는 시행 1의 Score 배열을 가져와서 비교한다. 기존 모델의 출력값과 Score 배열을 비교하면 (B 0.8 > A 0.55)는 만족하지만 (B 0.8 > C 0.9)는 위배되므로 B의 Score 0.8을 C의 Score 0.9보다 큰 값을 가지도록 업데이트한다. 다음 시행도 반복적으로 수행하여 Score 값이 업데이트되지 않을 때까지 반복한다. Fig. 21.는 Spectral MLE의 3단계의 과정을 거쳐 업데이트가 진행되지 않을 때까지 수행 후

시행	A	B	C	D	선호도	시행 결과
1 (seed)	0.55	0.8	0.9	0.65	A < B, C	A Score 변경하지 않음
2	0.5	<b>0.95</b>	0.9	0.65	B > A, C	B Score > 0.9 이상으로 변경
3	0.5	0.95	0.9	0.65	B > C > A, D	C Score 변경하지 않음
4	0.5	0.95	0.9	0.65	D < B, C	D Score 변경하지 않음
최종	0.5	0.95	0.9	0.65	B > C > D > A	

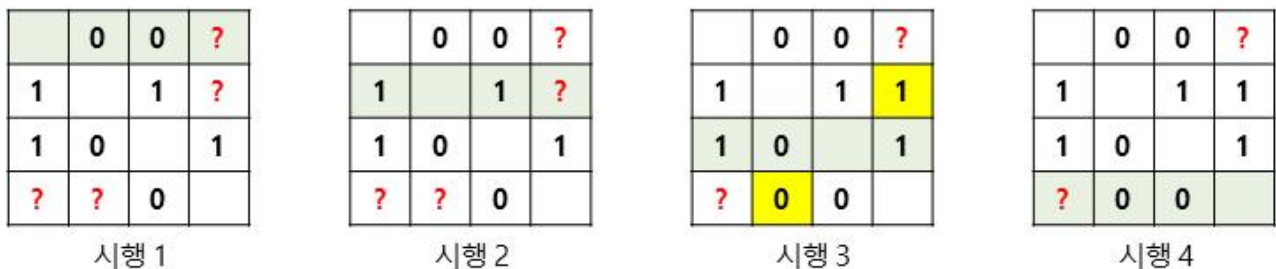


Fig. 21. [Step 3] Example of Refine Process

Score를 정렬한 결과이다. 최종적으로  $B > C > D > A$  와 같은 Listwise 순위와 점수로 변환할 수 있다.

앞서 정의한 9개의 토픽별 시험용 데이터셋을 학습이 완료된 모델에 입력하여 Pair 형태의 값들을 출력하고, Spectral MLE 후처리 방법을 이용하여 Listwise 출력값으로 도출된 결과를 가지고 성능 평가를 진행한다.

### 4.3 Evaluation Metrics for Ranking problems

일반적으로 추천 시스템에서의 추천 순위와 관련된 데이터의 검증 및 성능 평가를 위해 MAP, nDCG를 측정한다[14]. MAP(Mean Average Precision)는 정답 순위가 모델의 출력값 순위와 ‘맞다(0)/아니다(1)’의 정밀도(Precision)를 평균하여 계산한 값이다[15]. nDCG(normalize Discounted Cumulative Gain)는 추천된 순위와 연관성 점수를 고려하며 분모에 로그(log)를 사용하여 순위가 높을수록 가중치를 높게 측정하여 성능을 평가하는 지표이다[16]. MAP는 실제 정답 순위와 모델이 나온 결과에 대해 ‘순위가 맞다/아니다’ 이항 분류로 판단하기 때문에 1위와 2위의 순위 관계성, 2위와 5위에 대한 순위 관계성에 대해서는 고려하지 않기 때문에 실험에 대한 성능 지표는 nDCG를 가지고 성능을 평가하고 검증한다.

nDCG는 수식 (3)과 (4)를 활용한 수식 (5)로 표현할 수 있다. 수식 (3)의 DCG는 순위가 낮을수록 평가지표에 미치는 영향을 줄여서 관련성 점수를 합한 값을 의미하며  $rel_i$ 는  $i$ 번째 순위에 해당하는 종합점수이다. 수식 (4)의 IDCG는 이상적인 상황에서의 DCG 값을 의미하며 모든 순위가 정확하게 맞은 값을 의미한다. 수식 (5)의 nDCG는 DCG 값을 IDCG 값으로 나누어 DCG 값을 정규화한 값을 의미한다. 각 DCG와 IDCG는 분모에 log를 사용함으로 우선에 가까운 정답 순위를 학습모델이 순위를 맞췄을 때 가중치를 크게 부여한다. nDCG가 1에 가까울수록 정답 순위와 모델 결과 순위가 대부분 예측했다는 것을 의미한다.

$$DCG = \sum_{i=1}^k \frac{rel_i}{\log_2(i+1)} \quad (3)$$

$$IDCG = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (4)$$

$$nDCG = \frac{DCG}{IDCG} \quad (5)$$

### 4.4 Experimental Result

학습모델은 토픽별 시험용 데이터셋에 대해 평가하며 최종 성능은 토픽별 성능 결과를 평균해서 구한다. 선행연

구 [6]는 Pairwise 비교 모델의 실험 결과에 관해 서술하였다. 모델의 실험 결과가 전체 순위를 도출하는 것이 2개의 조합가설이 한 쌍(a, b)이 입력되었을 때, 단순 조합가설 a와 조합가설 b 중 어느 것이 우선인지 결과로 도출하고, Spectral MLE 알고리즘을 활용한 전체 조합가설들의 순위를 선별하는 방안을 제안하였다.

본 연구는 기존 선행연구 [6]를 확장하여 Fig. 8.의 조합가설 a, b, c, d, e의 조합가설 쌍 10개의 각 조합가설 쌍의 이진 분류(a, b 중 a가 우선)가 아닌 후처리 방법을 이용한 결과인 조합가설 a~e의 모든 1~5위까지의 순위를 도출하였다. Table 11.은 후처리 알고리즘을 이용하여 총 9개의 토픽에 대해 nDCG 성능을 평가한 실험 결과로 토픽 전체의 nDCG 평균이 0.8654로 계산되는 것을 확인할 수 있다. 전장상황과 유사한 상황분석의 데이터를 생성하기 위해 모의 시나리오 보고서 기반 시스템에 직접 레이블을 작업하고 자동화한 데이터셋을 구축하였다. 학습모델을 통해 구축한 데이터셋을 nDCG 성능 지표로 검증한 결과, 데이터셋이 비교적 올바르게 생성됨을 확인하였다.

Table 11. Experimental Result using nDCG metric

Topic	nDCG Score
#1	0.8816
#2	0.8515
#3	0.8720
#4	0.8764
#5	0.8627
#6	0.8555
#7	0.8648
#8	0.8592
#9	0.8655
Avg.	0.8655

nDCG가 높을수록 상위권 순위에 해당하는 순위들을 모델이 제대로 추론했다고 볼 수 있다. 지휘관에게 정확하고 신뢰도가 높은 정보들을 주는 것은 중요하지만 가장 우선인 1위의 순위만 가지고 의사결정을 하는 것보다 여러 개의 높은 순위들의 결과들을 가지고 의사결정 지원하는 것이 중요하므로 F-measure, MRR, MAP 등 다른 지표보다 nDCG의 평가지표를 가지고 의사결정을 판단하는 것을 추천한다.

## V. Conclusion

본 논문에서는 기존의 선행연구 [4]인 ‘모의 시나리오 기반 가설 데이터 셋 생성 방법’의 다음 단계의 가설 데이터셋 구축 방안 연구로 지휘관의 의사결정지원 및 미래 전장인식을 위해 최종적인 전장상황분석 결과에 필요한 데이터셋 생성 방안을 제안하였다. 또한, 조합가설 데이터셋 생성 방안에 대해 수동으로 사람이 직접 레이블할 수 있도록 학습 데이터셋 생성도구 SW 시스템을 구성하였고, 자동으로 시스템에서 자동화 알고리즘을 통해 레이블 하는 방안에 대해 제안하였다. 생성한 데이터셋은 Siamese Network 모델을 통해 쌍 형태의 출력값을 도출하고, 전체 순위를 도출할 수 있는 Spectral MLE 알고리즘을 적용하여 nDCG의 성능 지표로 검증하였다.

본 논문은 실 전장상황에 관련된 데이터를 활용한 것이 아닌 실 데이터와 유사한 모의 시나리오를 설계하였고, 웹에서 쉽게 취득할 수 있는 뉴스 보도자료, 국방백서, 국방일보 등의 데이터를 활용하여 데이터 구축에 활용하였다. 실 데이터가 아닌 유사한 데이터를 설계하여 직접적으로 군에서 사용하는 체계인 지능형 지휘통제체계에 활용하기에는 적합하지 않다. 사람이 직접 수동으로 데이터셋을 레이블 작업을 할 수 있지만 좀 더 각 속성의 중요도를 판별하여 자동화한 데이터 레이블 작업 생성 방안에 관해 기술하였기 때문에 추후 인공지능 기반 지능형 지휘통제체계의 학습모델에 필요한 데이터셋 생성 시 제안한 방법을 활용할 수 있다.

자동화 알고리즘에서 조합가설 데이터셋의 속성 중 관련성 지표가 중요하지만 본 논문에서는 관련성 생성 방안에 대한 설계 내용은 기술하지 않았다. 관련성 생성 방안 연구와 관련성과 연관이 있는 Fig. 4의 첫 번째 과정에서 필요한 후보가설과 관련된 사용자 질의 데이터셋 구축 생성 방안, 검증모델로 표현한 회귀 모델 및 머신러닝 앙상블을 비교한 연구 등은 추후 논문을 통해 제시할 예정이다. 또한, 국방품질 측면에서 직접 구축한 데이터셋을 지능형 지휘통제체계에 활용할 수 있도록 지속적으로 보완하여 품질을 향상할 예정이다.

## ACKNOWLEDGEMENT

This work was supported by the Defense Acquisition Program Administration and the Defense Rapid Acquisition Technology Research Institute (UC200018D).

## REFERENCES

- [1] C.E. Lee, J.H. Son, H.S. Park, S.Y. Lee, S.J. Park, Y.T. Lee, “Technical Trends of AI Military Staff to Support Decision-Making of Commanders,” *Electronics and Telecommunications Trends*, Vol. 36, No. 1, pp. 89-98, Feb. 2021, doi:10.22648/ETRI.2021.J.360110.
- [2] F. Barlos et al. “Collection and Monitoring via Planning for Active Situational Scenarios (COMPASS): Strategic Multi-Layer Assessment Report,” . Sandia National Laboratories, Jan. 2020, doi:10.2172/1592839.
- [3] DEFENSE ADVANCED RESEARCH PROJECTS AGENCY, Active Interpretation of Disparate Alternatives, <https://www.darpa.mil/program/active-interpretation-of-disparate-alternatives>
- [4] E. Cho, S. Jin, Y. Shin, and W. Lee, “A Virtual Battlefield Situation Dataset Generation for Battlefield Analysis based on Artificial Intelligence,” *Journal of the Korea Society of Computer and Information*, vol. 27, no. 6, pp. 33-42, Jun. 2022, doi:10.9708/JKSCI.2022.27.06.033.
- [5] H. Jo, H. Choo, S. Jin, Y. Shin, W. Lee, K. Shin, “Hierarchical Graph Neural Network for Identifying Relevance between Hypotheses that Describe Battlefield Situation”, *Database Research*, vol. 38, no. 2, pp. 18-28, Dec. 2022.
- [6] G. Suh, Y. Shin, S. Jin, W. Lee, J. Ahn, and C. Suh, “Recommendation Model for Battlefield Analysis based on Siamese Network,” *Journal of the Korea Society of Computer and Information*, vol. 28, no. 1, pp. 1-8, Jan. 2023, doi:10.9708/JKSCI.2023.28.01.001.
- [7] J. Chen, P. Hu, E. Jimenez-Ruiz, O. M. Holter, D. Antonyrajah, and I. Horrocks, “OWL2Vec\*: embedding of OWL ontologies,” *Machine Learning*, vol. 110, no. 7, pp. 1813-1845, Jun. 2021, doi:10.1007/s10994-021-05997-6.
- [8] J. Pennington, R. Socher, and C. Manning, “Glove: Global Vectors for Word Representation,” *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, doi: 10.3115/v1/d14-1162.
- [9] Y. Shin, S. Jin, and J. Ahn, “A study on embedding method of hierarchical dataset for AI based battlefield situation analysis”, *Korea Software Congress 2022*, vol. 49, no. 2, pp. 81-83, Dec. 2022.
- [10] M. Köppel, A. Segner, M. Wagener, L. Pensel, A. Karwath, and



- S. Kramer, "Pairwise Learning to Rank by Neural Networks Revisited: Reconstruction, Theoretical Analysis and Practical Performance," *Lecture Notes in Computer Science*, pp. 237–252, 2020, doi:10.1007/978-3-030-46133-1\_15.
- [11] N. Serrano and A. Bellogín, "Siamese neural networks in recommendation," *Neural Computing and Applications*, vol. 35, no. 19, pp. 13941–13953, May. 2023, doi:10.1007/s00521-023-08610-0.
- [12] Y. Chen, and C. Suh, "Spectral mle: Top-k rank aggregation from pairwise comparisons," *International Conference on Machine Learning*, PMLR, pp. 371–380, Jun. 2015, doi:10.48550/arXiv.1504.07218
- [13] S. Negahban, S. Oh, and D. Shah, "Rank Centrality: Ranking from Pairwise Comparisons," *Operations Research*, vol. 65, no. 1, pp. 266–287, Feb. 2017, doi:10.1287/opre.2016.1534.
- [14] M. Chen, "Performance Evaluation of Recommender Systems," *International Journal of Performability Engineering*, vol. 13, no. 8, pp. 1246–1256, Dec. 2017, doi: 10.23940/ijpe.17.08.p7.12461256.
- [15] A. Moffat and J. Zobel, "Rank-biased precision for measurement of retrieval effectiveness," *ACM Transactions on Information Systems*, vol. 27, no. 1, pp. 1–27, Dec. 2008, doi: 10.1145/1416950.1416952.
- [16] K. Järvelin and J. Kekäläinen, "IR evaluation methods for retrieving highly relevant documents," *ACM SIGIR Forum*, vol. 51, no. 2, pp. 243–250, Aug. 2017, doi:10.1145/3130348.3130374.

## Authors



Yukyung Shin received the B.S. degree in Mathematics, Computer Science from Duksung Women's University, Korea, in 2017, and received the M.S. degree in Data Science from Ajou University, Korea, in 2019.

Yukyung Shin is currently an engineer in Hanwha Systems. She is interested in Deep Learning, Recommender System, Military Communications and Network.



Soyeon Jin received the B.S. degree in Computer Engineering from Chonbuk National University, Korea, in 2003. Soyeon Jin is currently a chief engineer in Hanwha Systems. She is interested in data links,

machine learning, military communications, and unmanned systems.



Jongchul Ahn received the B.S. degree in Electronic Engineering from Inha University, Korea, in 1997, and received the M.S. degree in NCW Engineering from Ajou University, Korea, in 2017.

Jongchul Ahn is currently a chief engineer in Hanwha Systems. He is interested in SW development, satellite communication, NCW Engineering and military communications.