

분류번호 :	2001070305_23v2
능력단위 명칭 :	인공지능 데이터 특징 추출
능력단위 정의 :	인공지능 데이터 특징 추출이란 모델의 성능을 높이기 위하여 도메인 지식을 활용하여 객체의 특징을 발견하고 학습에 필요한 주요 특징을 선택하는 능력이다.

능력 단위 요소	수행 준거
2001070305_23v2.1 인공지능 특징 발견하기	<p>1.1 도메인 지식을 바탕으로 확보된 데이터의 전체 내용을 파악할 수 있다. 1.2 파악된 데이터에서 추출할 대상 특징을 발견할 수 있다. 1.3 발견된 데이터의 특징과 유사한 인공지능 특징을 모아 분류할 수 있다. 1.4 분류된 인공지능 특징과 유형을 문서화할 수 있다.</p> <p>【지식】</p> <ul style="list-style-type: none"> • 인공지능 객체의 표현방식 • 신호처리 관련 지식 • 데이터 시각화 지식 • 초자동화(Hyperautomation) 지식 <p>【기술】</p> <ul style="list-style-type: none"> • 객체의 표현 방식별 입출력 기술 능력 • 인공지능 객체 분류 능력 • 시각화 프로그램 활용능력 <p>【태도】</p> <ul style="list-style-type: none"> • 전체 특징을 입체적으로 파악하려는 노력 • 새로운 기술의 특징을 적극적으로 파악하려는 노력
2001070305_23v2.2 인공지능 특징 생성하기	<p>2.1 발견된 특징을 생성하기 위한 알고리즘을 선정할 수 있다. 2.2 선정된 알고리즘이 가질 수 있는 값의 범위와 형태를 정의할 수 있다. 2.3 정의된 데이터를 바탕으로 특징 추출 알고리즘의 입출력을 설계할 수 있다. 2.4 설계된 알고리즘을 적용하여 결정된 특징을 생성할 수 있다. 2.5 생성된 특징을 문서화할 수 있다.</p> <p>【지식】</p> <ul style="list-style-type: none"> • 인공지능 객체의 표현방식 • 인공지능 특징의 추상화 표현 방법론 • 연속형/범주형 데이터 지식 • 초자동화(Hyperautomation) 지식 <p>【기술】</p> <ul style="list-style-type: none"> • 특징의 변수화 처리 능력 • 벡터화 표현 능력 • 인공지능 객체 분류 능력 <p>【태도】</p> <ul style="list-style-type: none"> • 이전에 대해 적극적으로 해결하는 자세 • 특징 생성 기술에 대해 지속적으로 탐구하려는 노력
2001070305_23v2.3 인공지능 특징 결정하기	<p>3.1 생성된 특징들을 벡터 형태로 변환할 수 있다. 3.2 변환된 특징 벡터를 분류할 수 있다. 3.3 도메인 지식을 기초로 하여 분류된 특징의 중요도와 영향력을 평가할 수 있다. 3.4 평가 결과에 따라 분류된 특징 중에 가장 중요한 특징을 선택할 수 있다. 3.5 선택된 특징들을 척도화하여 표현 방법을 결정할 수 있다.</p>

	<p>【지식】</p> <ul style="list-style-type: none"> • 인공지능 주요 특징의 결정방법 • 인공지능 주요 특징의 중요도와 영향력 • 음성 및 언어 기반 지능
2001070305_23v2.3 인공지능 특징 결정하기	<p>【기술】</p> <ul style="list-style-type: none"> • 탐색적 데이터 분석(EDA) 능력 • AI기반 사이버 보안 적용 능력 • 차원 축소 기법
	<p>【태도】</p> <ul style="list-style-type: none"> • 핵심 특징을 정확히 파악하려는 노력 • 새로운 특징의 결정 기술에 대해 지속적으로 탐구하려는 자세
	<p>4.1 선택된 특징을 결합하거나 분할하여 발견되지 않았던 새로운 특징을 생성할 수 있다.</p> <p>4.2 생성된 새로운 특징을 척도화하여 수치로 표현할 수 있다.</p> <p>4.3 새롭게 생성된 특징 중 학습에 필요한 중요성을 재판단할 수 있다.</p>
2001070305_23v2.4 인공지능 특징 개선하기	<p>【지식】</p> <ul style="list-style-type: none"> • 특징에 대한 중요도 • 입력데이터의 특징 • 윤리적이고 설명 가능한 AI <p>【기술】</p> <ul style="list-style-type: none"> • 인공지능 알고리즘 구현 능력 • 대형 언어 모델(LLM) 활용 능력 • MLOps 활용 능력 <p>【태도】</p> <ul style="list-style-type: none"> • 핵심 특징을 개선하려는 적극적인 자세 • 새로운 특징 기술을 지속적으로 탐구하려는 노력

□ 적용범위 및 작업상황

고려사항

-이 능력단위는 모델의 성능을 높이기 위하여 도메인 지식을 활용하여 객체와 특징을 발견하고 학습에 필요한 주요 특징을 선택하는 업무에 적용한다.

-객체는 실제 적용대상에서 의미 있는 실체를 말한다.

-모델의 성능이란 정확도(Accuracy), 효율성(Efficiency), 일반화 능력(Generalization)을 말한다.

-인공지능 주요 특징의 결정방법 지식에는 LDA(Linear Discriminant Analysis), PCA(Principal Component Analysis), SVD(Singular Value Decomposition), ICA(Independent Component Analysis) 기법 등을 포함한다

-이 능력단위는 객관적인 지표를 선정하여 최적의 인공지능 학습에 필요한 주요 특징을 도출하는 것이 목표이다.

-알고리즘은 모델을 학습시키기 위한 수학적인 공식과 절차를 말한다. 이 알고리즘은 데이터를 처리하고 분석하여 패턴을 찾아내고, 예측 및 분류를 수행하는 등의 작업을 수행한다.

-인공지능 모델의 특징을 벡터화한다는 것은 입력 데이터의 특성이나 속성을 모델에서 쉽게 처리할 수 있는 숫자 형식, 특히 벡터로 변환하는 것을 말한다. 간단히 말해서 입력 데이터를 일련의 숫자로 변환하여 AI 모델에서 더 이해하기 쉽고 사용할 수 있도록 만드는 프로세스이다.

-주요특징이라 하면 인공지능 학습에서 불필요한 연산을 줄이고 학습결과의 정확도를 높이기 위해 데이터에 내재된 특징 중 학습 전에 주요한 특징들을 미리 선정한 것을 말한다.

-명세화된 결과는 인공지능 문제정의 능력단위에서 정의된 것을 기반으로 한다.

-주요특징 및 주요특징의 선정 결과는 객관적으로 비교할 수 있어야 한다.

-인공지능 모델링에서는 입력 데이터에서 가장 중요한 변수나 속성을 선택하는 과정인 특징 추출에 다양한 알고리즘을 사용할 수 있다.

-주성분 분석(PCA)은 원본 데이터 세트를 주성분이라고 하는 상관관계가 없는 새로운 변수 세트로 변환하는 차원 축소 기술을 말한다. 이러한 구성 요소는 원본 데이터 세트의 대부분의 정보를 유지하면서 변수 수를 줄인다.

-선형 판별 분석(LDA)은 데이터 세트에서 서로 다른 클래스를 가장 잘 구분하는 기능의 선형 조합을 찾는 것을 목표로 하는지도 학습 방법을 말한다. LDA는 특히 분류 문제에 유용하다.

-t-SNE(t-Distributed Stochastic Neighbor Embedding)는 특히 고차원 데이터 세트를 시각화하는 데 적합한 비선형 차원 감소 기술을 말한다. 원래 공간과 저차원 공간의 확률 분포 사이의 차이를 최소화하여 데이터의 로컬 구조를 보존하는 것을 목표로 한다.

-오토인코더는 입력 데이터를 압축하고 재구성하는 방법을 배우는 감독되지 않은 신경망을 말한다. 오토인코더의 숨겨진 레이어는 입력 데이터의 특징 표현 역할을 할 수 있으며 분류 또는 클러스터링과 같은 다양한 작업에 사용할 수 있다.

-LLM(Large Language Model, 대형 언어 모델)이란, 대량의 텍스트 데이터로 훈련된 인공지능 모델을 말한다.

-MLOps (Machinelearing Operations)란, 인공지능 모델의 개발과 운영을 위한 프로세스와 방법론을 말한다.

-인공지능 FATE란 인공지능의 공정성(Fairness), 책임성(Accountability), 투명성(Transparency), 윤리의식(Ethics)을 의미한다.

-인공지능 윤리기준이란 3대 기본원칙에 따라 인간성(Humanity)을 구현하기 위해 인공지능의 개발 및 활용 과정에서 인간의 존엄성 원칙, 사회의 공공선 원칙, 기술의 합목적성 원칙을 지키는 것을 의미한다.

자료 및 관련 서류

- 인공지능 데이터 객체 정의서
- 인공지능 데이터 특징 정의서

- 인공지능 데이터 특징 결정서

장비 및 도구

- 컴퓨터, 인터넷, 문서작성 프로그램
- 데이터 분석용 소프트웨어 혹은 패키지
- 분석 데이터 처리용 프로그래밍 언어 실행환경
- 데이터 시각화 프로그램

재료

- 해당사항 없음

□ 평가지침

권장평가방법

- 평가자는 능력단위 인공지능 데이터 특정 추출의 수행준거에 제시되어 있는 내용을 평가하기 위해 이론과 실기를 나누어 평가하거나 종합적인 결과물의 평가 등 다양한 평가 방법을 사용할 수 있다.
- 평가자는 다음 사항을 평가해야 한다.

권 장 평 가 방 법	평 가 유 형	
	과 정 평 가	결 과 평 가
A.포트폴리오		
B.문제해결 시나리오		V
C.서술형시험	V	V
D.논술형시험	V	V
E.사례연구	V	V
F.평가자 질문		
G.평가자 체크리스트	V	V
H.피평가자 체크리스트		
I.일지/저널		
J.역할연기		
K.구두발표	V	V
L.작업장평가		
M.기타		

평가시 고려사항

- 수행준거에 제시되어 있는 내용을 성공적으로 수행할 수 있는지를 평가해야 한다.
- 평가자는 다음 사항을 평가해야 한다.
 - 인공지능 모델을 위한 특징을 발견하는 능력
 - 발견된 특징을 생성하는 능력
 - 생성한 인공지능 특징을 선택하는 능력
 - 선정된 인공지능 특징을 개선하는 능력

□ 관련기초능력

순번	관련기초능력	
	주요영역	하위영역
1	의사소통능력	경청 능력, 기초외국어 능력, 문서이해 능력, 문서작성 능력, 의사표현 능력
2	수리능력	기초연산 능력, 기초통계 능력, 도표분석 능력, 도표작성 능력
3	문제해결능력	문제처리 능력, 사고력
4	정보능력	정보처리 능력, 컴퓨터활용 능력
5	기술능력	기술선택 능력, 기술이해 능력, 기술적용 능력

□ 개발·개선 이력

구 분	내 용	
직무명칭(능력단위명)	인공지능모델링(인공지능 데이터 특징 추출)	
분류번호	기준	2001070305_19v1
	현재	2001070305_23v2
개발·개선연도	현재	2023
	최초(1차)	2019
버전번호	v2	
개발·개선기관	현재	
	최초(1차)	정보기술·사업관리 인적자원개발위원회(한국IT비즈니스진흥협회)
향후 보완 연도(예정)	-	