

Microsoft Azure AI-900(인공지능)

01. Azure에서 AI 시작

키워드: “AI”

- AI란 무엇인가?
 - 정의: 인간의 행동과 기능을 모방하는 소프트웨어를 만드는 것
 - 주요 워크로드
 1. 기계 학습: AI 시스템의 기초, 데이터를 통해서 예측하고 결론을 도출하기 위한 컴퓨터 모델을 가르치는 방법
 2. 변칙 검색: 시스템에서 오류 또는 비정상적인 활동을 자동으로 감지할 수 있는 기능
 3. 컴퓨터 비전: 카메라, 비디오 및 이미지를 통해 세계를 시각적으로 해석할 수 있는 소프트웨어 기능
 4. 자연어처리: 컴퓨터가 글씨 또는 음성 언어를 해석하고 동일하게 응답할 수 있는 기능
 5. 지식 마이닝: 종종 대용량의 비정형 데이터에서 정보를 추출하여 검색 가능한 지식 저장소를 만드는 기능
- 기계 학습 작동 방법
 1. 야생화 샘플 데이터 수집
 2. 샘플을 종류별로 분류
 3. 종류 사이의 관계 알고리즘 탐색
 4. 알고리즘 결과 모델에 요약
 5. 새로운 샘플 발견 시 올바른 식별 가능
- Microsoft Azure의 Machine Learning
 - 기능 및 특징
 1. 자동화된 기계학습
 2. Azure Machine Learning Designer
 3. 데이터 및 컴퓨팅 관리
 4. 파이프라인
- 변칙 검색 이해
 1. 차량 센서의 데이터 수집
 2. 수집한 데이터를 토대로 예상 범위 학습
 3. 측정값이 예상 범위 벗어날 시 엔지니어에게 알림
 4. 급유 정차 시 레이저에게 문제 해결할 수 있도록 도움
- 컴퓨터 비전 이해
 - 정의: 시각적 처리를 다루는 AI 분야
 - 종류
 1. 이미지 분류: 콘텐츠에 따라 이미지를 분류하는 기계 학습 모델 교육
 2. 개체 감지: 이미지 내의 개별 개체를 분류하고 경계 상자를 사용하여 해당 위치를 식별
 3. 의미 체계 구분: 이미지를 픽셀 단위로 정교하게 분석
 4. 이미지 분석: 이미지에서 정보를 추출하는 고급 이미지 분석 기술 + 기계 학습 모델
 5. 얼굴감지, 분석 및 인식: 이미지에서 사람의 얼굴을 찾는 특수한 형태의 개체 감지
 6. OCR(광학 문자 인식): 이미지의 텍스트를 검색하고 읽는 데 사용되는 기술
- Microsoft Azure의 Computer Vision
 - 기능과 특징
 1. Computer Vision: 이미지 및 영상 분석 설명, 태그, 개체 및 텍스트 추출
 2. Custom Vision: 사용자 고유 이미지를 사용하여 이미지 분류 및 개체 검색
 3. 얼굴: Face 서비스를 사용하여 얼굴 감지 및 얼굴 인식
 4. From Recognizer: 특정 양식의 문서나 청구서에서 정보 추출

- 자연어 처리 이해

- 정의: 문어와 구어를 이해하는 소프트웨어 생성을 다루는 AI 영역

1. 음성 언어 해석: 소리에서 글자 그대로를 가져옴
2. 언어 간 구어, 문어구 자동 번역
3. 텍스트 분석 및 해석
4. 명령 해석 및 적절한 작업 결정
5. 음성 응답 합성

- Microsoft의 자연어 처리

1. 언어: 텍스트 이해 및 분석, 언어 모델 학습, 지능형 애플리케이션 빌드
2. Translator: 텍스트를 60개 이상의 언어로 번역
3. Speech: 음성을 인식, 음성 합성, 음성 언어 번역
4. Azure Bot: 에이전트와의 대화 기능, 채팅 봇 만들기 및 관리

- 지식 마이닝 이해

- 정의: 대량의 비정형 데이터에서 정보 추출 후 지식 저장소 생성

- Microsoft Azure의 지식 마이닝

- Azure Cognitive Search: 데이터를 모아서 색인을 매긴 후에 저장 -> 내부, 외부에서 쉽게 검색할 수 있도록 해줌
- 이미지 처리, 자연어 처리와 같은 Azure의 AI 기능을 사용하여 더 놀라운 검색 가능

- 담당 AI의 이해

- AI는 6가지 원칙에 의해 설계

1. 공정성
2. 신뢰성 및 안정성
3. 개인정보 보호 및 보안
4. 포용성
5. 투명성
6. 책임성

02-1. 자동화된 기계학습 사용

키워드: “기계 학습, Azure Machine Learning, Azure 자동화된 Machine Learning, AutoML 프로세스 이해”

- 기계 학습

- 수학 및 통계를 사용하여 알 수 없는 값을 예측할 수 있는 모델을 만드는 기술
- x = 기능, y = 예측된 레이블, f = 계산하는 함수

- Azure Machine Learning Studio

- 태스크를 간소화하며, 데이터를 준비하고, 모델을 학습 시키고, 예측 서비스를 배포하는 클라우드 기반 서비스
- 작업영역: 데이터, 컴퓨팅 리소스, 코드, 모델 및 기계 학습 워크로드와 관련된 기타 아티팩트 관리
- 컴퓨팅: 데이터 과학 활동에 대한 컴퓨팅 대상을 관리
- 컴퓨팅 인스턴스: 데이터 및 모델 작업
- 컴퓨팅 클러스터: 실험 코드 처리를 지원하는 가상 머신
- 유추 클러스터: 예측 서비스의 배포 대상
- 연결된 컴퓨팅: 기존 Azure 컴퓨팅 리소스에 연결

- Azure 자동화된 Machine Learning

- 자동화된 기계 학습 기능 포함: 최고 성능의 감독 모드 기계 학습 모델을 찾음
- 프로그래밍 지식 없이도 모델 학습 가능
- 알고리즘 선택 및 하이퍼 매개변수 튜닝을 자동화

- AutoML 프로세스 이해
- 데이터 준비: 필요에 따라 데이터 사전 처리 또는 정리 및 변환
- 모델 학습: 데이터를 학습 및 유효성 검사 집합을 두 그룹으로 분할
 학습 데이터 집합: 기계 학습 모델 학습
 유효성 검사 데이터 집합: 기계 학습 모델에서 성능 테스트
- 성능 평가: 모델의 예측이 알려진 레이블과 얼마나 가까운지 비교
- 예측 서비스 배포: 다른 사용자가 사용할 수 있도록 서버 또는 디바이스에 애플리케이션으로 모델 배포

02-2. ML디자이너 사용 회귀모델 만들기

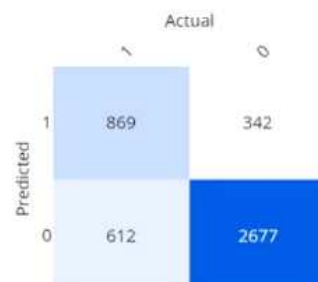
키워드: “회귀 기계 학습 시나리오 식별, Azure Machine Designer, 회귀 단계 이해”

- 회귀 기계 학습 시나리오 식별
- 회귀: 원하는 결과를 예측하기 위해 변수 간의 관계를 이해하는 데 사용되는 기계학습의 한 형태
- 모델이 학습을 통해 특징 조합을 레이블에 맞추도록 레이블에 대한 특징과 알려진 값을 모두 포함하는 데이터를 사용하여 모델을 학습하는 지도 기계 학습 기법
- Azure Machine Learning Designer
- 기계 학습 모델을 학습, 테스트 및 배포하는 데 사용할 수 있는 디자이너라는 시각적 인터페이스
- 성능 평가
- MAE(평균 절대 오차): 예측 값과 실제 값 사이의 평균 차이
- RMSE(평균 제곱 오차): 예측 값과 실제 값 사이의 평균 제곱 파이의 제곱근
- RSE(상대 제곱 오차): 차이에 대한 제곱을 기준, 0~1사이의 상대 메트릭
- RAE(상대 절대 오차): 예측 값과 실제 값 간의 절대 차이에 대한 제곱을 기준, 0~1 사이의 상대 메트릭
- 결정 계수(R2): 예측 값과 실제 값 간의 분산이 어느 정도인지 개략적으로 알려줌. 1에 가까울수록 좋은 것

02-3. ML디자이너 사용 분류모델 만들기

키워드: “분류 기계 학습 시나리오 식별, Azure Machine Designer, 분류 단계 이해”

- 분류 기계 학습 시나리오 식별
- 분류: 항목이 속한 범주 또는 클래스를 예측하는 데 사용
- 레이블에 대한 특징과 알려진 값을 모두 포함하는 데이터를 사용하여 모델을 학습하는 감독형 기계 학습 기술
 ex) 임상 데이터를 사용하여 환자가 아플지 여부 예측, 기록 데이터를 사용하여 은행이 대출을 제공해야 하는지 여부 예측
- 혼동 행렬
- 예측 값과 실제 값이 모두 1인 경우: 진양성
- 예측 값과 실제 값이 모두 0인 경우: 진음성
- 정확도: 올바른 예측(진양성 + 진음성)과 총 예측 수의 비율
- 정밀도: 올바르게 식별된 양성 사례의 비율
- 재현율: 양성으로 분류된 사례 중 실제로 양성인 비율
- F1-Score: 정밀도와 재현율의 결합



02-4. ML디자이너 사용 클러스터링 모델 만들기

키워드: “Azure Machine Learning 작업 영역 만들기, 컴퓨팅 리소스 만들기, 데이터 탐색, 학습 파이프라인 만들기 및 실행, 클러스터링 모델 평가, 유추 파이프라인 만들기, 예측 서비스 배포”

- 클러스터링
- 유사한 항목을 특성에 따라 클러스터로 그룹화 하는 데 사용되는 기계 학습의 한 형태
- 감독되지 않은 기계 학습의 한 예로, 해당 특성 또는 기능에 따라 항목을 클러스터로 분리하는 모델 학습
- Microsoft Azure Machine Learning 디자이너를 사용하여 코드를 작성할 필요 없이 끌어다 놓기 시각적 인터페이스를 사용하여 클러스터링 모델 제작 가능

03-1. CV 사용하여 이미지 분석

키워드: “Azure에서 이미지 분석”

- 소개
 - 컴퓨터 비전: AI 애플리케이션이 세상을 보고 이해할 수 있는 솔루션을 만드는 데 중점을 둠
 - 용도
 1. 콘텐츠 조직: 사람의 사람 또는 개체를 식별, ID에 따라 구성
 2. 텍스트 추출: 텍스트가 포함된 이미지와 PDF 문서를 분석하고 텍스트를 구조적 형식으로 추출
 3. 공간 분석: 공간에서 개체 식별 및 해당 공간 내에서 이동을 매핑
- Computer Vision용 Azure 리소스
 - Cognitive Services: Text Analytics 및 Translator Text 등의 다른 많은 인지 서비스와 함께 Computer Vision을 포함하는 일반적인 인지 서비스 리소스
- 이미지 설명
 - 이미지를 분석하고 감지된 개체를 평가하고 구문이나 문장을 생성하여 이미지에서 감지된 내용을 설명
- 시각적 특징 태그 지정
 - Computer Vision에서 생성되는 이미지 설명은 이미지에 대한 태그를 제안하는 데 사용
- 개체 감지
 - 경계 상자 좌표를 반환 -> 개체의 형식 및 좌표 집합도 제공
- 이미지 분류
 - 콘텐츠를 기준으로 이미지를 분류
 - 부모/자식 계층 구조 사용, 감지된 개체가 기존 범주와 비교되어 분류를 제공하는 최적의 방법 결정

03-2. CV 사용하여 이미지 분류

키워드: “분류 이해, Azure에서 이미지 분류”

- 이미지 분류
 - AI 애플리케이션의 일반적인 작업
 - 기계 학습의 예측 능력을 사용하여 AI시스템이 이미지에 기반한 실제 항목을 식별할 수 있도록 함
 - ex) 제품 식별, 재해 조사, 의료 진단
- 분류 이해
 - 범주 또는 클래스를 예측하기 위해 기계 학습 분류 기법 사용
 - 분류의 대상으로 되는 개체가 사진과 같은 이미지인 기계 학습의 한 가지 기법
- Custom Vision용 Azure 리소스
 - Custom Vision: 학습 리소스, 예측 리소스
 - Cognitive Services: Custom Vision + 인지 서비스

03-3. CV 사용하여 이미지 개체 감지

키워드: “분류 이해, Azure에서 이미지 개체 감지”

- 개체 감지
 - 이미지에서 개체의 개별 유형을 인식한 다음 해당 개체의 위치를 식별하도록 모델 학습
 - ex) 건물 안전 검사, 주행 지원, 종양 감지
- 개체 검색의 개념
 - 이미지에서 식별된 각 개체의 클래스, 개체 분류의 확률 점수, 각 개체의 경계 상자 좌표

03-4. Face 서비스를 이용해 얼굴 감지 및 분석

키워드: “Azure의 Face 분석”

- 얼굴 감지 및 분석
 - 알고리즘을 사용하여 이미지 또는 비디오 콘텐츠에서 사람의 얼굴을 찾고 분석하는 AI 영역
 - 얼굴 감지: 얼굴 주위에 사각형을 형성하는 경계 상자 좌표 반환
 - 얼굴 인식: 얼굴 특징에서 알려진 개인을 식별
 - 얼굴 분석: 얼굴 랜드마크(눈, 코, 눈썹, 입술 등)와 같은 다른 정보를 반환
- 얼굴 감지 및 분석의 용도
 - 보안, 소셜 미디어, 지능형 모니터링, 광고, 실종자, 신원 확인
- Azure의 Face 분석 시작하기
 - Computer Vision: 이미지 주위에 경계 상자 좌표를 반환하는 등 얼굴 감지 및 몇 가지 기본 얼굴 분석
 - Video Indexer: 비디오에서 얼굴을 감지 및 식별
 - Face: 얼굴 감지, 인식 및 분석

03-5. CV 서비스로 텍스트 읽기

키워드: “Azure에서 읽기 API”

- OCR(광학 문자 인식)
 - 개별 모양을 문자, 숫자, 문장 부호 또는 텍스트의 다른 요소로 인식하도록 모델을 학습
- ex) 필기, 의료 기록이나 기록 문서 디지털화, 은행 예금을 위해 인쇄
- 읽기 API
 - 텍스트가 많은 스캔된 문서 처리
 - 인쇄된 텍스트가 있는 이미지 지원, 필기를 인식하여 사용할 적절한 인식 모델을 자동으로 결정하는 기능
1. API에 이미지를 제출하고 응답으로 작업 ID 검색
 2. 작업 ID를 사용하여 이미지 분석 작업의 상태를 확인
 3. 작업의 결과 탐색

03-6. Form Recognizer 이용 영수증 분석

키워드: “Azure에서 영수증 분석”

- Form Recognizer Service
 - OCR을 사용하여 양식의 필드를 디지털화하여 문제 해결
- Azure에서 영수증 분석 시작
 - 양식 데이터를 해석할 수 있는 예측 모델 + 최첨단 OCR
 - 필드 이름을 값에 일치
 - 데이터 테이블 처리
 - 날짜, 전화번호, 주소, 합계 등과 같은 특정 유형의 필드 식별

04-1. 언어 서비스 사용하여 텍스트 분석

키워드: “텍스트 분석”

- 텍스트 분석
 - 텍스트 내용에 대한 인사이트를 얻기 위해 문서 또는 구의 다양한 측면 평가
 - AI 알고리즘이 텍스트에서 동일한 특성을 평가하여 구체적인 인사이트를 알아내는 프로세스
- 언어 감지 서비스
 - 텍스트에서 지배적 언어에 중점을 둠
 - 언어에 대한 구 길이 또는 총 텍스트 양과 같은 지배적 언어를 결정

- 엔터티 인식

- 언어 서비스에 구조화되지 않은 텍스트를 입력하면 인식한 텍스트로 엔터티의 목록을 반환

04-2. 음성 인식 및 합성

키워드: “Azure 음성 인식”

- 음성 인식 및 합성

- 음성 인식: 음성 입력을 감지하고 해석하는 기능

ex) 녹화된 동영상 또는 라이브 비디오에 대한 자막 제공, 전화 통화 또는 회의 내용 대본 만들기, 자동화된 메모 받아쓰기, 추가 처리를 위해 의도한 사용자 입력 결정

- 음성 합성: 음성 출력을 생성하는 기능

ex) 사용자 입력에 대한 음성 응답 생성, 전화 시스템용 음성 메뉴 만들기, 기차역 또는 공항과 같은 공공장소에서 공지 사항 발송

04-3. 텍스트 및 음성 번역

키워드: “Azure에서 번역”

- 직역 및 의미론적 번역

- 번역 서비스 제공 시 의미론적 맥락을 고려해 영어 번역을 반환해야 함

- 고려 대상: 문법 규칙(공식/비공식) 및 구어체

- 텍스트 및 음성 번역

- 텍스트 번역: 한 언어에서 다른 언어로 문서를 번역

- 음성 번역: 음성 언어 간 번역, 곧바로 번역 또는 중간 과정으로 텍스트 형식을 거쳐 번역

- Azure에서 번역

- 텍스트 간 번역을 지원하는 Translator 서비스

- 음성 텍스트 변환 및 음성 간 번역을 가능하게 하는 Speech 서비스

- Speech 서비스를 통한 음성 번역

- Speech-to-text: 오디오 원본에서 텍스트 형식으로 음성을 전사하는 데 사용

- Text-to-Speech: 텍스트 소스에서 음성 오디오를 생성

- Speech Translation: 한 언어의 음성을 다른 언어의 텍스트 또는 음성으로 변환하는 데 사용

04-4. CLU 사용 언어모델 만들기

키워드: “Conversational Language Understanding”

- CLU

- 샘플 발화와 다른 종류의 입력 값이 들어와도 의도 및 엔터티를 예측하도록 학습 가능

- 사전 빌드된 일반적인 시나리오의 엔터티(목적어), 의도 컬렉션 제공, 모델의 시작점으로 사용할 수 있으며 사용자 고유의 엔터티(목적어) 및 의도 데이터를 만들 수 있음

04-5. 언어 서비스 및 Azure Bot 서비스 사용하여 봇 빌드

키워드: “언어 서비스 및 Azure Bot Service 시작”

- 질문과 답변

- Language Studio의 사용자 지정 질문 답변 기능을 사용하여 질문 및 답변 쌍으로 구성된 기술 자료를 만들 수 있음

- Azure Bot Service를 사용하여 봇 빌드

- Microsoft Bot Framework SDK를 사용하여 대화 흐름을 제어, 기술 자료와 통합하는 코드 작성

05. Anomaly Detection 소개

키워드: “Anomaly Detection 정의, 작동 방식, 사용해야 하는 경우”

- 변칙 검색
 - 계열의 값이 예상 매개 변수 내에 있는지 여부를 결정하는 데 사용되는 AI 기술
- Anomaly Detection
 - 변칙: 예상 값 또는 값 범위를 벗어난 값
 - 측정값이 음영이 지정된 경계를 벗어나면 주황색 점을 사용하여 값이 비정상적으로 간주됨을 표시
 - 정확한 변칙 검색은 신속한 문제 해결로 이어져 수익 손실을 방지하고 브랜드 평판을 유지하는 데 도움이 됨
- Anomaly Detection 작동 방식
 - Anomaly Detector 서비스가 변칙을 식별 -> 경계는 민감도 값을 사용하여 설정
 - 상위 경계 및 하위 경계는 expectedValues, upperMargin 및 lowerMargin이라는 개념을 사용하여 계산
 - $upperBoundary = expectedValue + (100 - marginScale) * upperMargin$
- Anomaly Detection을 사용해야 하는 경우
 - 일괄 검색: 한 번에 전체 데이터 계열에 알고리즘을 적용하는 작업
 - 데이터 집합을 일괄 처리로 평가하는 작업 -> 시계열을 사용하여 변칙 검색
 - 실시간 탐지

06. 지식 마이닝

키워드: “Azure Cognitive Search 검색 솔루션의 요소, 기술 세트를 사용하여 보강 파이프라인 정의, 인덱스”

- Azure Cognitive Search
 - 다양한 구조적, 반구조적 및 비구조적 문서에서 데이터를 추출하는 검색 솔루션을 만들기 위한 인프라와 도구 제공
 - 기능
 1. 모든 원본의 데이터
 2. 전체 텍스트 검색 및 분석
 3. AI 기반 검색
 4. 다국어
 5. 지역 지원
 6. 구성할 수 있는 사용자 환경
- 검색 솔루션의 요소 확인
 - 데이터 아티팩트가 포함된 데이터 원본으로 시작 -> Azure Storage 폴더 및 파일의 계층 구조, Azure SQL Database, Azure Cosmos DB와 같은 데이터베이스의 텍스트 -> 네이티브 형식으로 자동화
- 기술 세트를 사용하여 보강 파이프라인 정의
 - 전체 텍스트 검색을 위해 인덱싱할 수 없는 콘텐츠에서 텍스트 및 정보를 추출
 - 기술 세트는 검색 가능하도록 데이터를 추출하고 보강하는 작업을 정의
- 인덱스 이해
 - Azure Cognitive Search 인덱스 = 검색할 수 있는 문서의 컨테이너
 - 특성: 디자인할 때 필드에서 필요한 동작을 설정하는 것을 잊어버린 경우 해당 기능을 얻는 유일한 방법
 - 테이블 프로젝트: 쿼리 및 시각화를 위해 관계형 스키마에서 추출된 데이터를 구조화
 - 개체 프로젝트: 데이터 엔티티를 나타내는 JSON 문서
 - 파일 프로젝트: 이미지를 JPG 형식으로 저장하는 데 사용