

멀티캠퍼스 혁신성장 청년인재 집중양성

인공지능 자연어처리(NLP) 기반 기업 데이터 분석 과정

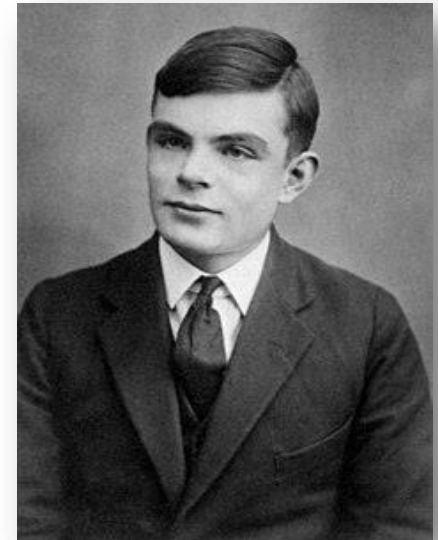


02. 인공지능(AI)의 시작

인공지능(AI)의 시작

자연어 처리에 대한 관심은 1950년 앨런 튜링(Alan Turing)이 이른바 튜링 테스트(Turing Test)가 등장한 "Computing Machinery and Intelligence"라는 논문을 발표하면서 본격적으로 시작되었다.

*"인간이 컴퓨터와 대화하고 있다는 것을 깨닫지 못하고
인간과 대화를 계속할 수 있다면
컴퓨터는 지능적(Intelligence)인 것으로 간주될 수 있다."
- 앨런 튜링 -*



앨런 매티슨 튜링은 영국의 수학자, 암호학자, 논리학자이자 컴퓨터 과학의 선구적 인물이다. 알고리즘과 계산 개념을 튜링 기계라는 추상 모델을 통해 형식화함으로써 컴퓨터 과학의 발전에 지대한 공헌을 했다. 튜링 테스트의 고안으로도 유명하다.

03. 인공지능 시장과 활용사례

| 알파고는 '지도학습'과 '강화학습'의 강점을 결합한 기계학습 기법을 사용해 바둑을 '열공'하고, 2016년 3월 이세돌 9단의 대국에 등장했다. 이세돌 9단이 이길 거라는 대다수의 예상을 깨고 알파고는 4대1로 대승을 거뒀다. 이 사건은 사람들에게 충격과 함께 인공지능에 새롭게 관심을 가지는 계기를 주었다.

| 인공지능 알고리즘을 이용해 사용자와 음성으로 의사소통하는 AI 스피커가 속속 등장하고 있다. 현재 외국 AI 스피커로는 아마존 '에코'를 시작으로 구글의 '구글 홈', 애플 '홈팟', 마이크로소프트 '인보크'가 있다. 국내에서는 SKT의 '누구', KT '기가지니', 네이버 '웨이브', 그리고 최근 합류한 카카오의 '카카오미니'가 있다.



[바둑기사]

※출처: Huffpost, 60전 전승 알파고, 이세돌의 역사적 선언
https://www.huffingtonpost.kr/Arthur-jung/story_b_14094264.html



[인공지능 스피커]

※출처: 조선비즈, "주인말만 듣는다" 더 똑똑해진 AI 스피커
http://biz.chosun.com/site/data/html_dir/2018/05/03/2018050303715.html

03. 인공지능 시장과 활용사례

- | 기타큐슈시 대형 택시업체 다이이치교통산업은 일본 손해보험사 및 글로벌 컨설팅 기업 액센츄어와 공동으로 졸음 예측 공동 연구에 인공지능을 사용한다. 2017년 3월부터 5월 약 두 달 간 택시 100대와 택시기사 100명을 대상으로 택시 운전자의 심박수, 운전자 태도, 운전자의 주행 모습 등의 데이터를 수집해 인공지능을 학습했다.
- | 국내 인공지능 스타트업 마인드셋이 개발한 인공지능 CCTV '마인드아이'는 실시간으로 영상을 측정하고 원하는 상황·물체를 확인한다. 컴퓨터와 스마트폰 카메라로 사용할 수 있어 별도의 하드웨어가 필요 없다. 기존 CCTV보다 사람·사물을 찾는 시간을 절약할 수 있게 도와준다. 많은 인파 속 특정 인물을 찾거나 미아를 찾는 데 유용하게 활용할 수 있다.



[졸음운전 방지]

※출처:오토데일리, 운전 중 졸음 인공지능으로 완벽대응, 파나소식, 첨단 제어시스템 개발
<http://www.autodaily.co.kr/news/articleView.html?idxno=336657>



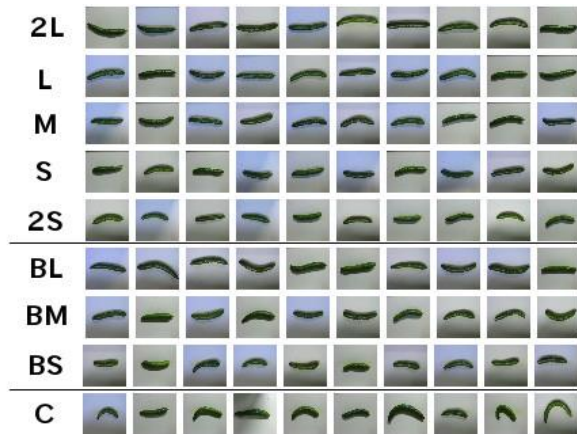
[실시간 영상 속 인물 인식]

※출처:블로터, [블로터11th] 알아두면 쓸데있는 신기한 인공지능 50선
<https://www.bloter.net/archives/289626>

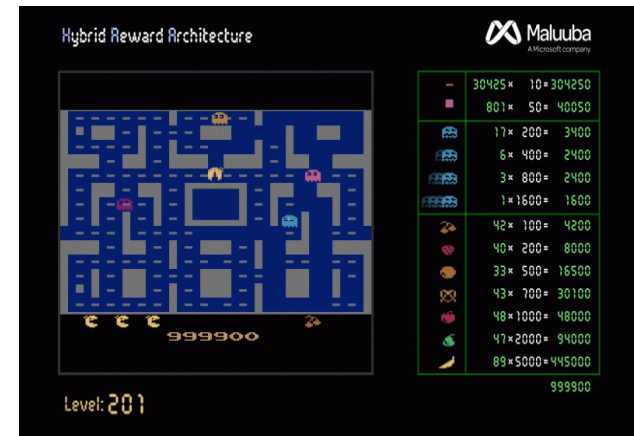
03. 인공지능 시장과 활용사례

| 자동차 임베디드 시스템 디자이너였던 마코토 코이케 씨는 일을 그만두고 부모님을 도와 오이농장 일을 시작했다. 오이에 정성을 다 쏟아부어도 힘든데, 그보다는 오이를 9등급으로 분류하는 일이 손도 많이 가고 시간도 더 소요되었다. 알파고와 이세돌 9단의 대국에서 아이디어를 얻은 마코토 씨는 딥러닝 기반의 오이 자동분류 시스템을 만들었다.

| '팩맨' 게임은 '유령'을 피해 도망 다니면서 '먹이'를 다 먹으면 이긴다. 지금까지 최고점은 266,330점이었다. 그런데 그 기록을 깬 이가 등장했다. 사람이 아니다. 마이크로소프트 인공지능 시스템 '말루바'는 999,999점으로 팩맨 게임에서 만점을 기록했다.



[오이 분류]



[아케이드 게이머]

※출처:Google Blog, How a Japanese cucumber farmer is using deep learning and TensorFlow
<https://cloud.google.com/blog/products/gcp/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow>

※출처:블로터,블로터11월알아두면 쓸데있는 신기한 인공지능50선
<https://www.bloter.net/archives/289626>

04. 인공지능 자연어처리(NLP)란?

자연어 처리란?

자연어 처리(自然語處理) 또는 자연 언어 처리(自然言語處理)는 인간의 언어 현상을 컴퓨터와 같은 기계를 이용해서 모사할 수 있도록 연구하고 이를 구현하는 인공지능의 주요 분야 중 하나다.

정보 검색, QA 시스템, 문서 자동 분류, 신문기사 클러스터링, 대화형 Agent 등 다양한 응용이 이루어지고 있다.

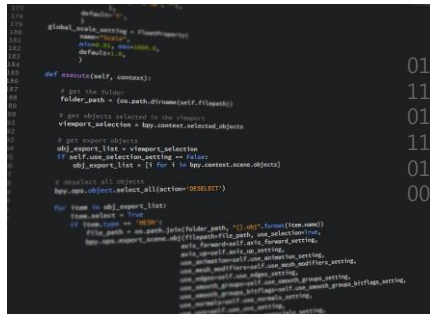
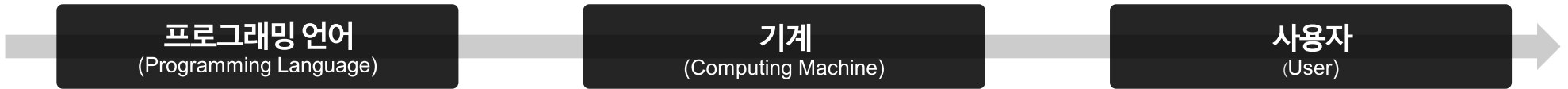


※ 출처: https://ko.wikipedia.org/wiki/%EC%9E%90%EC%97%B0%EC%96%B4_%EC%B2%98%EB%A6%AC

04. 인공지능 자연어처리(NLP)란?

전통적인 프로그래밍 언어

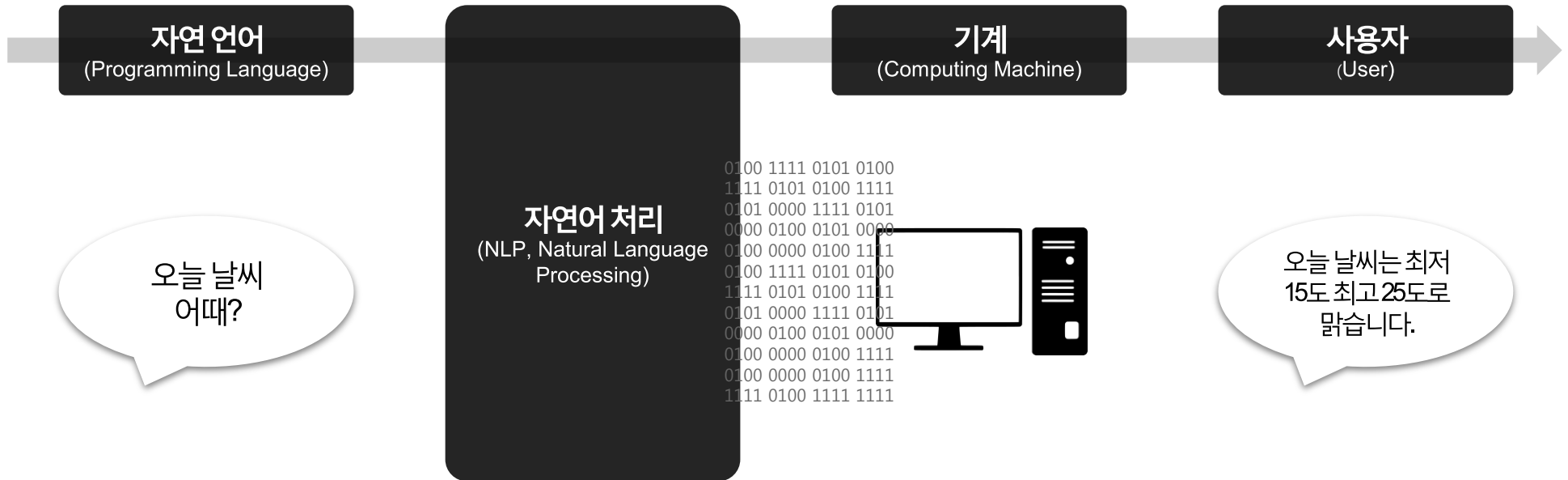
: 기계(혹은 컴퓨터)를 실행하기 위해서 기계가 이해할 수 있는 프로그래밍 언어로 명령을 내리고,
그 결과를 사용자에게 전달



04. 인공지능 자연어처리(NLP)란?

자연어 처리

: 인간의 언어(=자연 언어)로 명령을 내리면 기계가 자연어 처리(NLP)를 통해 이해하여 처리하고,
그 결과를 사용자에게 전달



04. 인공지능 자연어처리(NLP)란?

자연어 처리란?



전통적인 프로그래밍 언어가 인간이 기계 언어로 기계(=컴퓨터)를 이해시키는 것이었다면,
자연어 처리는 기계가 인간의 언어(=자연 언어)를 이해하여 소통하는 것을 말한다.

04. 인공지능 자연어처리(NLP)란?

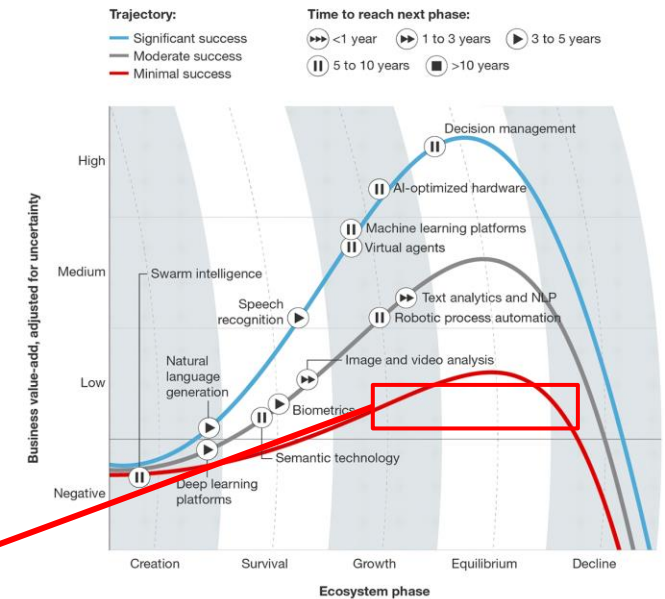
자연어 처리 기술 전망

- 인공지능 기술 내에서도 자연어 처리는 빠르게 성장
- 기술전문 매체 테그레이더(TechRadar) 자료를 보면 인공지능 기술 중에서도 가능, 빠르게 성장하는 기술

FORRESTER RESEARCH

TechRadar™: Artificial Intelligence Technologies, Q1 '17

TechRadar™: Artificial Intelligence Technologies, Q1 2017



▶▶ Text analytics and NLP

※ 출처: Forbes, Top 10 Hot Artificial Intelligence (AI) Technologies

<https://www.forbes.com/sites/gilpress/2017/01/23/top-10-hot-artificial-intelligence-ai-technologies/#4b97ce791928>

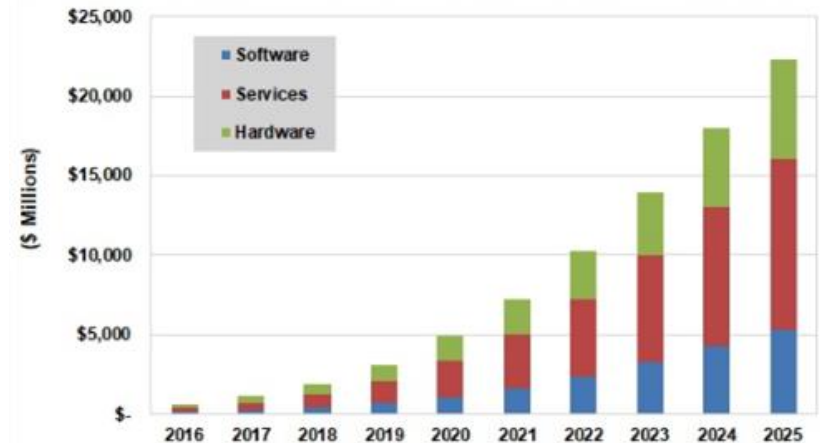
04. 인공지능 자연어처리(NLP)란?

자연어 처리 시장 전망

- 2016년 \$500M(한화 5,600억원)에서
- 2025년 \$22.3B(한화 24.9조)로 증가(10년 내 44.6배 성장)
- 자연어 처리 시장 성장 동력은 "수요 증가"
- 인공지능 스피커와 같은 스마트 장치 사용 증가
- 웹 및 클라우드 기반 비즈니스 응용프로그램 증가
- 비정형 데이터(Unstructured data)로부터 인사이트를 도출 Needs 증가



Natural Language Processing Total Revenue by Segment, World Markets: 2016-2025



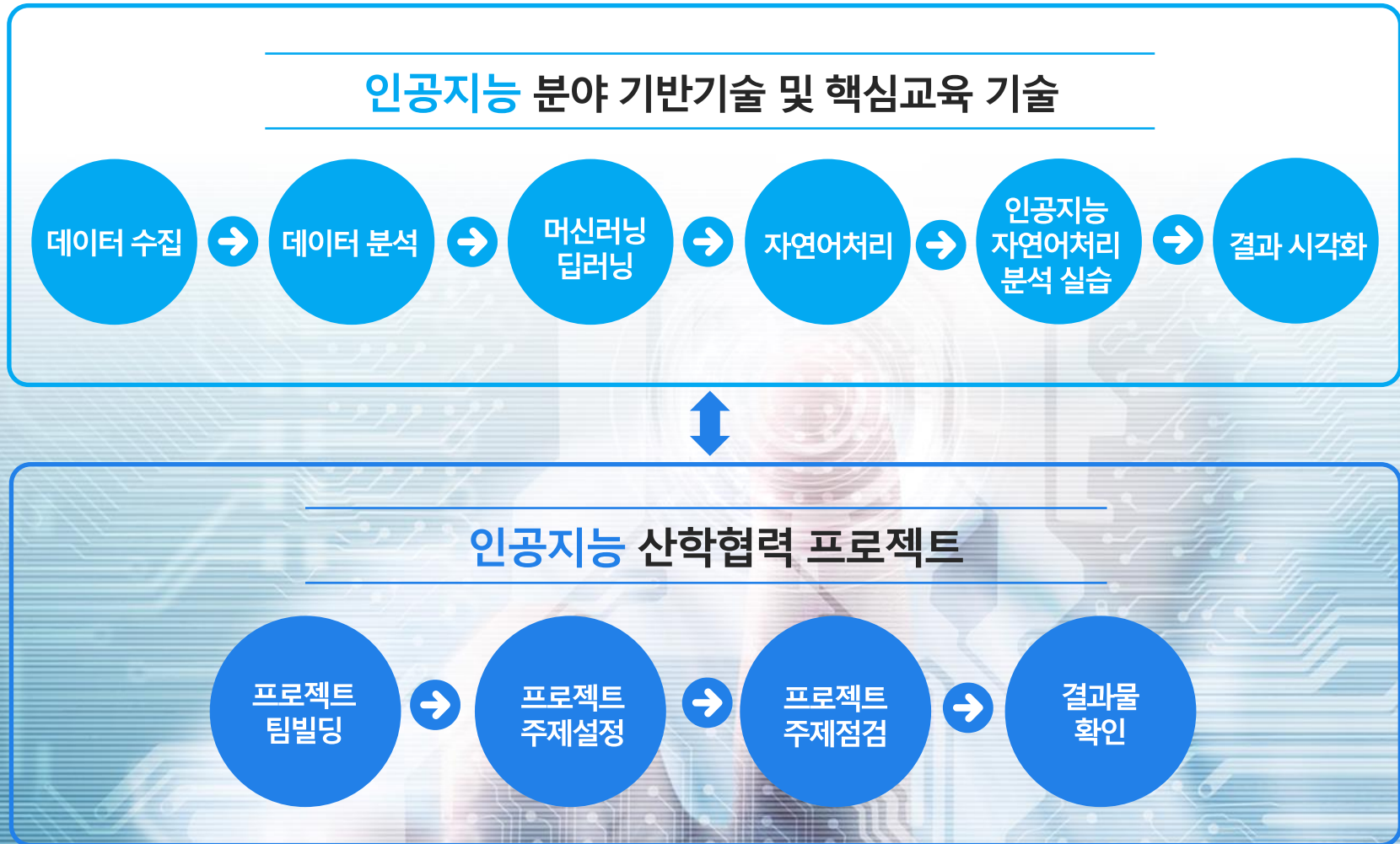
Source: Tractica

※ 출처: Tractica, Natural Language Processing Market to Reach \$22.3 Billion by 2025

<https://www.tractica.com/newsroom/press-releases/natural-language-processing-market-to-reach-22-3-billion-by-2025/>

05. 과정개요

- | 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여 현장 실무 기반의 교육 환경 구축
- | 파이썬 프로그래밍, 데이터 분석, 자연어처리 분석 관련 활용 기술 학습을 통해 인공지능 자연어처리 기업데이터 분석 전문가 양성



06. 과정소개

- | 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여 현장 실무 기반의 교육 환경 구축
- | 파이썬 프로그래밍, 데이터 분석, 자연어처리 분석 관련 활용 기술 학습을 통해 인공지능 자연어처리 기업데이터 분석 전문가 양성

과정명

인공지능 자연어처리(NLP)기반
기업 데이터 분석전문가

교육일정

2019.05.20(월) ~ 2019.11.08(금)

1. 학습안내



교육 기간

• 총 6개월, 920시간 (8시간/일)

교육 시간

• 09:00 ~ 18:00

수강 정원

• 48명

06. 과정소개

2. 학습 대상

- 인공지능 관련 분석 직무로 **취업**을 원하는 분
- 자연어처리(NLP)기반 기업 데이터 분석에 대한 **실무 프로젝트까지** 교육을 원하는 분

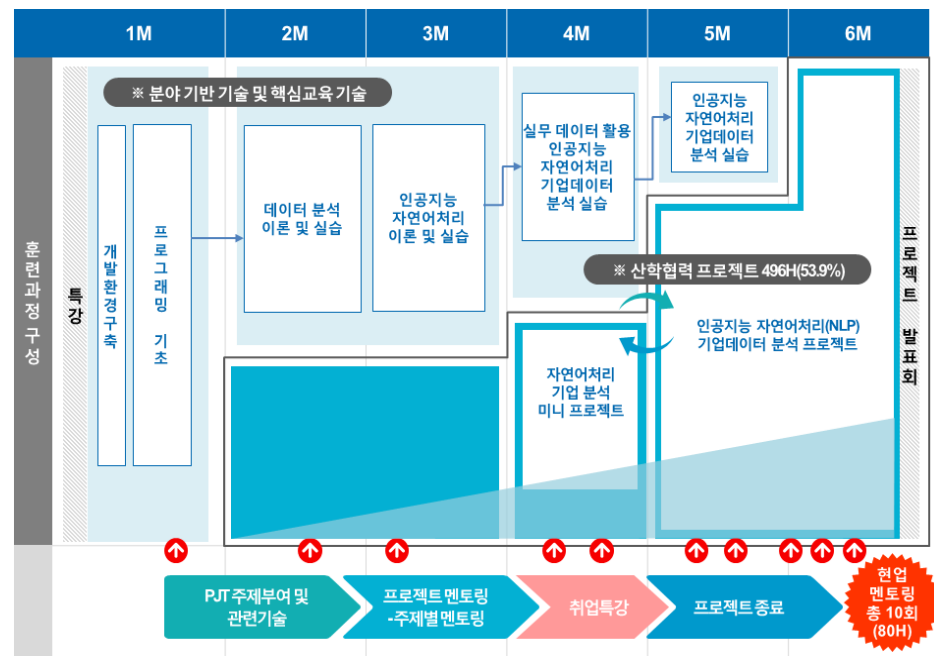
3. 학습 강점

- 협력기관으로부터 도출된 **실제 수요를 기반**으로 커리큘럼 구성
- 인공지능 자연어처리 기업데이터 분석에 필요한 **데이터 분석 능력**과 **인공지능 자연어 처리 분석 능력**을 모두 함양할 수 있도록 설계
- 베테랑 전문가가 검증한 일관성 있는 교육흐름을 통해 실무에 필요한 기술을 갖춘 **맞춤형 인재로 성장**
- 실무 기업데이터를 사용한 **인공지능 자연어처리 기업데이터 분석실습**을 통해 분석기술 체득
- **과정별 전담 매니저를 배치**하여 훈련 시작 전, 훈련 중, 훈련 종료, 사후관리까지 체계화된 운영/관리 프로세스를 갖추고 있으며, 이에 관련한 모든 행정 제반 절차를 보유함
- 4차 산업혁명시대의 업계 대응 전략과 최근 동향 그리고 IT 비즈니스에 대한 미래 진화 방향과 대응전략을 수립하는 IT 트렌드 특강 진행

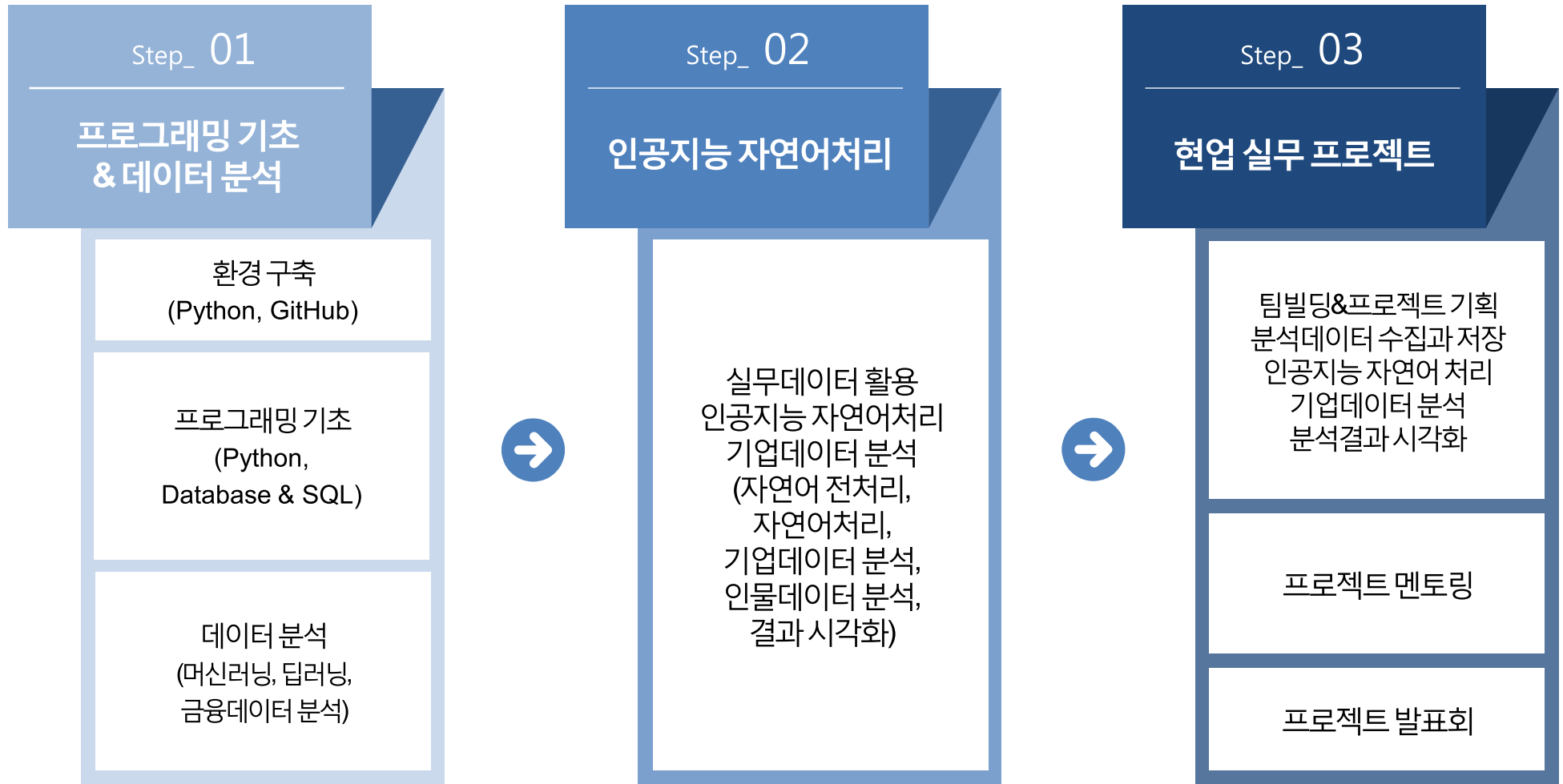
06. 과정소개

4. 학습 진행

- 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여, 현장 실무 기반 교육 환경 구축
- 기초부터 데이터 수집, 데이터 분석, 분석 결과 시각화까지 학습하고
인공지능 자연어처리 기업데이터 분석을 실습함
- **대용량(1천만 건) 실무 기업데이터**로 인공지능 자연어처리 기업데이터 분석 실습
- 인공지능 자연어처리 기업데이터 분석에 필요한
데이터 분석 능력과 인공지능 자연어처리 분석 능력 함양
- 협력기관의 PJT 과제 도출 후,
주제선정 3단계를 거쳐 산학협력 과제 선정
- 인공지능/IT/금융/핀테크의 관련 직무 혹은 서비스를 하고 있는
현업 전문가가 실무 프로젝트 멘토로 참여하여,
각 직무에 맞는 데이터의 이해 및 활용방법에 대한 노하우를
프로젝트 내 데이터 수집, 분석, 시각화 등에 적용 할 수 있도록
지도/조언하고 완성도 높은 산출물을 낼 수 있도록 멘토링함.



07. 과정 이수 체계도



IT트렌드 특강

08. 상세 커리큘럼(1/3)

- | 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여 현장 실무 기반의 교육 환경 구축
- | 파이썬 프로그래밍, 데이터 분석, 자연어처리 분석 관련 활용 기술 학습을 통해 인공지능 자연어처리 기업데이터 분석 전문가 양성

단계	교과목명	단원명	세부내용	교육시간
특강	오리엔테이션 및 특강	• IT 트렌드 특강		8시간
공통	개발환경 구축	• 개발환경 구축 • 소스 공유 및 관리	• 파이썬 개발환경 설정 • 파이썬 기본 문법 • Git을 사용한 소스 공유 및 관리	8시간
프로그래밍	프로그래밍 기초	• 파이썬 프로그래밍 • 웹크롤링을 활용한 데이터 수집 • RestAPI를 활용한 데이터 수집 • 데이터베이스 & SQL	• 웹크롤링 활용 데이터 수집, RestAPI 활용 데이터 수집 • 관계형 데이터베이스 설계 및 SQL 실습	112시간
데이터 분석	데이터 분석 이론 및 실습	• 머신러닝 이론 및 실습 • 딥러닝 이론 및 실습 • 금융데이터 분석 이론 및 실습	• 머신러닝 개요, 동향, 머신러닝 실습 • 딥러닝 개요, 동향, Neural networks, 딥러닝 실습 • 금융데이터 수집, 통계적 분석, 시계열 분석, 금융모델링	112시간

※ 커리큘럼은 교육 진도에 따라 변경 될 수 있습니다.

08. 상세 커리큘럼(2/3)

- | 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여 현장 실무 기반의 교육 환경 구축
- | 파이썬 프로그래밍, 데이터 분석, 자연어처리 분석 관련 활용 기술 학습을 통해 인공지능 자연어처리 기업데이터 분석 전문가 양성

단계	교과목명	단원명	세부내용	교육시간
인공지능 자연어 처리	인공지능 자연어처리 이론 및 실습	<ul style="list-style-type: none"> •영문 자연어 전처리 •한글 자연어 전처리 •자연어처리 비정형 텍스트 분석 	<ul style="list-style-type: none"> •영문 자연어 전처리 실습 •한글 자연어 전처리 실습 •개체명 인식, 주제어 찾기, 문서요약, 카테고리 분류 	104시간
기업데이터 분석	실무데이터 활용 인공지능 자연어처리 기업데이터 분석 실습	<ul style="list-style-type: none"> •분석 데이터 수집 및 정제 실습 •자연어 전처리 실습 •자연어 비정형 텍스트 분석 실습 •분석결과 시각화 •자연어처리 기업분석 개인 프로젝트 	<ul style="list-style-type: none"> •API를 활용한 뉴스데이터 수집 실습 •웹크롤링을 활용한 블로그 데이터 수집 실습 •뉴스, 블로그 데이터 전처리 실습 •해외기업/국내기업 관련 뉴스, 블로그 데이터 수집, 관련 브랜드 인지도 분석(예시, Netflix, 쿠팡) •해외인물/국내인물 관련 뉴스, 트위터 데이터 수집, 관련 인지도 분석(예시, Obama, BTS) •서비스(Yelp) 리뷰 데이터 수집, 시장반응 분석 •분석결과 시각화 (예시, 해외 상장기업 뉴스, 블로그, 소셜 데이터 수집 및 자연어처리 기업데이터 분석) 	112시간

※ 커리큘럼은 교육 진도에 따라 변경 될 수 있습니다.








08. 상세 커리큘럼(3/3)

- | 프로젝트와 기술교육을 연계한 **교육↔멘토링↔프로젝트**를 병행하여 현장 실무 기반의 교육 환경 구축
- | 파이썬 프로그래밍, 데이터 분석, 자연어처리 분석 관련 활용 기술 학습을 통해 인공지능 자연어처리 기업데이터 분석 전문가 양성

단계	교과목명	단원명	세부내용	교육시간
프로젝트	인공지능 자연어처리(NLP) 기업데이터 분석 프로젝트	• 팀 프로젝트	<ul style="list-style-type: none"> • 팀 빌딩과 프로젝트 기획 • 분석 데이터 수집과 저장 • 인공지능 자연어처리 기업데이터 분석 • 시각화를 위한 결과데이터 집계 • 분석 결과 시각화 • 분석 결과 리뷰 • 프로젝트 발표회 	384시간
멘토링	프로젝트 멘토링	• 프로젝트 멘토링	<ul style="list-style-type: none"> • 실무 프로젝트 개요 • 대체데이터 분석 적용 사례 소개 • 텍스트마이닝 적용 사례 소개 • 팀빌딩 & 프로젝트 기획 • 데이터 수집 및 저장 • 데이터 분석 • 결과 시각화 • 프로젝트 정리 	80시간

※ 커리큘럼은 교육 진도에 따라 변경 될 수 있습니다.

09. 활용장비

교과목명	활용 장비	사진	설명
개발환경 구축 프로그래밍 기초 데이터 분석 이론 및 실습 인공지능 자연어처리 이론 및 실습 실무데이터 활용 인공지능 자연어처리 기업데이터 분석 실습	Ubuntu		(Linux)운영체제
	Python Jupyter Notebook BeautifulSoup Pytorch	   	분석 도구 및 프레임워크
	MariaDB		데이터베이스
	GitHub		프로젝트 관리도구

※ 위에 소개된 장비들은 과정의 대표적인 활용 장비들로, 실제 과정에서는 이 외에도 다양한 프로그램들을 함께 활용합니다.

THANK YOU

Copyright by Multicampus CO.,LTD. All right reserved

www.multicampus.co.kr

서울특별시 강남구 언주로 508 10-17층

