

IT 기초 용어집

▣ 하드웨어

CPU

Memory

Disk

입출력장치

▣ 컴퓨터공학 · 과학

자료구조

알고리즘

운영체제

네트워크

데이터베이스

인공지능

가상화

리소스

블록체인

애자일 소프트웨어 개발 (애자일 방법론)

사물인터넷(IoT)

가상현실(VR)

증강현실(AR)

메타버스

소프트웨어 공학

임베디드 시스템

SI/SM

클라우드 컴퓨팅

▣ 프로그래밍

프로그래밍 언어

Python

Java

컴파일러

디버깅

객체 지향 프로그래밍

API

자연어처리

▣ 웹 개발

프론트엔드/백엔드

UI/UX

프레임워크

HTML

CSS

JavaScript

JSP

React

Node.js

Django

Java Spring

Vue.js

TypeScript

렌더링

컴포넌트

Servlet

Weka

▣ 데이터분석 · 시각화

- 공통 용어 -

빅데이터

EDA

전처리

시각화

머신러닝

데이터 마이닝

- 데이터베이스 관련 용어 -

RDBMS(관계형 데이터베이스)

NoSQL(Not only SQL)

SQL

스키마

데이터 모델링

Pandas

- 머신러닝 관련 용어 -

딥러닝

Supervised Learning (지도학습)

Unsupervised Learning
(비지도학습)

Classification(분류)

Regression(회귀)

Clustering(클러스터링)

Redis

NumPy

- 딥 러닝 관련 용어 -

CNN

RNN

LSTM

GRU

Transformer

▣ 그 외 주요 서비스

Git

GitHub

Notion

ChatGPT

Jira

AWS

Azure

GCP

SaaS

PaaS

IaaS

하드웨어

CPU central processing unit 의 약자로 컴퓨터 시스템을 통제하고 프로그램의 연산을 실행, 처리하는 가장 핵심적인 컴퓨터의 제어장치를 말합니다. 컴퓨터 프로그램의 명령어를 해석하여 연산하고, 외부로 출력하는 등 다양한 역할을 수행합니다.

Memory 컴퓨터 하드웨어와 관련해서 Memory 라는 단어를 사용할 때는 RAM 을 가리킵니다. RAM 은 CPU 가 처리하기 위한 데이터를 임시로 저장하는 장치입니다. 전원이 꺼지면 RAM 에 올라와 있던 데이터가 날아가는 휘발성이 특징입니다. 주기억 장치라고 부르기도 합니다.

Disk 컴퓨터 하드웨어와 관련해서 Disk 라는 단어를 사용할 때는 하드디스크 혹은 SSD 를 가리칩니다. 디스크는 대용량 데이터를 영속적으로 저장하는 장치입니다. 전원이 꺼지더라도 disk 에 저장된 데이터는 날아가지 않는 비휘발성 특징을 갖습니다. disk 에 저장된 데이터를 메모리에 로딩하고, CPU 로 처리하는 것이 일반적인 프로그램 동작 방식입니다. 보조기억 장치라고 부르기도 합니다.

입출력장치 컴퓨터에 입력 혹은 출력을 할 수 있도록 도와주는 장치들입니다. 대표적으로 키보드, **I/O Device** 마우스, 모니터, 프린터 등이 있습니다. 이들은 컴퓨터를 구성하는 필수적인 요소들은 아니기에 주변 장치라고도 불립니다.

컴퓨터공학 · 컴퓨터과학

자료구조
Data Structure 자료구조는 컴퓨터 프로그래밍에서 다루는 여러가지 형태의 데이터들을 더 효율적으로 처리하기 위해 특정한 형태로 저장하는 것을 말합니다. 예를들어 데이터를 들어온 순서대로 처리하는 큐, 나중에 들어온 데이터를 먼저 처리하는 스택, 노드와 엣지의 연결 형태로 표현하는 그래프와 부모, 자식 관계로 표현하는 트리 등이 있습니다. 이러한 자료구조는 복잡한 프로그램을 작성하는 기본 뼈대가 되고, 다양한 알고리즘들의 기반이 됩니다.

알고리즘
Algorithm 알고리즘은 주어진 문제를 논리적으로 해결하기 위해 필요한 절차, 방법, 명령어들을 모아놓은 것입니다. 넓게는 사람 손으로 해결하는 것, 컴퓨터로 해결하는 것, 수리적인 것, 비수학적인 것을 모두 포함합니다.

운영체제(OS)
Operating System 운영체제는 사용자가 하드웨어, 시스템 리소스를 제어하고, 프로그램에 대한 일반적인 서비스를 제공하는 소프트웨어입니다. 대표적인 예시로 윈도우, macOS, Linux 등이 있습니다.

네트워크
Network 컴퓨터들이 서로 연결되어 데이터를 교환하고 통신할 수 있는 구조를 말합니다. 이를 통해 정보를 빠르게 전달하고 자원을 공유할 수 있습니다. 인터넷은 수많은 네트워크의 집합체로 이루어져 있으며, TCP/IP 프로토콜을 기반으로 동작합니다.

데이터베이스
DataBase 구조화된 데이터의 집합을 관리하는 시스템입니다. 데이터베이스는 효율적인 데이터 관리를 위해 데이터를 조직화하고 검색, 갱신, 삭제하는 기능을 제공합니다. 관계형 데이터베이스(RDBMS)는 가장 흔히 사용되는 유형이며, MySQL, Oracle, PostgreSQL 등이 대표적입니다.

인공지능(AI)
Artificial Intelligence 컴퓨터 시스템이 인간과 비슷한 지능적인 작업을 수행하는 기술을 말합니다. 머신 러닝, 딥러닝, 자연어 처리 등이 인공지능의 하위 분야로 포함됩니다. 인공지능은 의료 진단, 자율 주행차, 언어 번역 등 다양한 분야에 응용됩니다.

가상화
Virtualization 하나의 물리적 자원을 여러 개의 가상 자원으로 분할하여 사용하는 기술입니다. 가상화를 통해 하나의 서버에서 여러 개의 가상 서버를 운영하거나, 가상 머신을 사용하여 여러 운영 체제를 동시에 실행할 수 있습니다. 이를 통해 자원 사용 효율을 높이고 유연성을 확보할 수 있습니다. 가상화 기술로는 VMware, Hyper-V, KVM 등이 있습니다.

리소스 사용자가 사용할 수 있는 컴퓨터의 하드웨어, 소프트웨어 자원을 가리킵니다. 연산을

많이 수행해야 하는 상황이면 CPU 리소스가 많이 필요하다, 메모리에 대용량의 데이터를 올려놓아야 하는 상황이면 메모리 리소스가 많이 필요하다, 통신을 통해서 데이터를 많이 다운로드 받아야하는 상황이면 네트워크 리소스가 많이 필요하다는 식으로 보통 이야기합니다.

블록체인 Block Chain	분산 데이터베이스 기술로, 여러 컴퓨터에 걸쳐 정보를 분산 저장하고 연결된 블록으로 구성합니다. 각 블록에는 이전 블록과 함께 암호화된 정보가 저장되어 무결성과 보안을 보장하며, 중앙 관리자 없이 모든 참여자가 데이터 변경을 검증할 수 있습니다. 이로써 데이터 위조나 변경이 어렵고, 신뢰성과 투명성을 제공하여 거래나 정보 교환을 보다 안전하게 할 수 있습니다.
애자일 소프트웨어 개발 Agile Software Development	주기로 프로젝트를 반복적으로 진행하고 지속적인 고객 피드백을 수용하여 유연하게 요구 사항을 조정하는 개발 방법론입니다. 초기 계획에 집착하는 대신 불확실성과 변화를 받아들이며, 팀원들은 스스로 조직하고 업무를 결정하여 효과적인 협력과 의사소통을 강조합니다. 이를 통해 고객의 요구를 빠르게 반영하고 높은 품질의 소프트웨어를 빠르게 제공하여 프로젝트를 성공적으로 완료합니다.
사물인터넷(IoT) Internet of Things	일상적인 사물들이 인터넷에 연결되어 데이터를 주고받고 상호작용하는 기술입니다. 센서, 카메라, 스마트 디바이스 등 다양한 사물들이 인터넷을 통해 정보를 수집하고 분석하여 현실 세계와 디지털 세계를 연결시키는 것이 특징입니다. 사물인터넷은 스마트 홈, 스마트 시티, 산업 자동화, 건강 관리 등 다양한 분야에서 활용되고 있으며, 우리의 생활과 사회를 더욱 편리하고 효율적으로 만드는데 기여하고 있습니다.
가상현실(VR) Virtual Reality	사용자를 컴퓨터 시뮬레이션된 가상 세계로 끌어들이는 인터랙티브한 체험 기술입니다. 헤드셋과 컨트롤러를 통해 사용자는 3D 환경에 몰입하여 마치 실제 세계에 있는 것처럼 느끼게 됩니다. 가상현실은 게임, 교육, 훈련, 의료 등 다양한 분야에서 활용되며, 사용자에게 새로운 경험과 감동을 제공하고 현실 세계에서는 어려운 상황을 안전하게 연습하거나 체험할 수 있는 기회를 제공합니다. 이러한 가상현실 기술은 더욱 진화하여 혁신적인 분야에 더 많은 가능성과 기회를 제시하고 있습니다.
증강현실(AR) Augmented Reality	실제 세계에 가상의 콘텐츠를 추가하여 혼합 현실을 제공하는 기술입니다. 스마트폰, 헤드업 디스플레이, AR 안경 등을 사용하여 실시간으로 실제 환경과 가상 콘텐츠를 결합시킵니다. 사용자는 주변 환경과 상호작용하면서 가상 객체를 볼 수 있고, 정보를 확인하거나 게임을 즐길 수 있습니다. 증강현실은 교육, 엔터테인먼트, 마케팅, 제조업 등 다양한 분야에서 활용되며, 사용자에게 새로운 경험과 가치를 제공하는 기술로 인기를 얻고 있습니다.

메타버스
Metaverse 현실 세계와 가상 세계가 융합된 공간으로, 인터넷 기반 가상 현실 환경을 의미합니다. 사용자는 가상 세계에서 자신을 대리인으로 나타내며, 다른 사용자와 상호작용하고, 가상 공간 내에서 활동하고 소통할 수 있습니다. 메타버스는 온라인 게임, 가상 쇼핑몰, 교육 플랫폼 등 다양한 분야에 적용되며, 더욱 현실과 유사하고 다양한 경험을 제공하는 디지털 세계로 주목받고 있습니다.

소프트웨어 공학 소프트웨어를 체계적으로 개발, 운용, 유지보수하는 학문과 기술을 의미합니다. 소프트웨어 개발 과정에서 요구사항 분석, 설계, 구현, 테스트, 배포, 유지보수 등 다양한 단계를 포함하며, 품질 향상과 효율성을 추구합니다. 소프트웨어 공학은 프로젝트 관리, 품질 보증, 소프트웨어 테스트 등에도 관련 기법과 절차를 적용하여 소프트웨어의 안정성과 신뢰성을 높이는데 초점을 둡니다. 소프트웨어 개발의 생명주기를 이해하고 효과적인 방법론과 도구를 활용하여 고품질의 소프트웨어를 개발하는데 사용됩니다.

임베디드 시스템
Embedded System 임베디드 시스템은 주로 특정 제어나 기능을 수행하는데 특화된 작고 자체적으로 동작하는 컴퓨터 시스템을 말합니다. 이러한 시스템은 주로 소형 디바이스, 전자제품, 자동차, 가전제품, 의료기기 등에 내장되어 사용됩니다. 실시간 처리, 저전력 운영, 안정성, 효율성이 필요한 분야에서 널리 활용되며, 별도의 사용자 인터페이스 없이 미리 정의된 기능을 자동으로 수행하는 것이 특징입니다.

SI
System Integration 시스템 통합 (SI) 다양한 하드웨어와 소프트웨어 구성 요소들을 하나로 통합하여 기능적으로 연결하고 원활한 동작을 보장하는 작업을 의미합니다. 주로 컴퓨터 시스템이나 정보 시스템을 구축하거나 업그레이드할 때 사용됩니다.

SM
System Management 시스템 관리 (SM): 시스템의 운영과 유지 보수를 효율적으로 수행하기 위한 활동을 말합니다. 시스템의 안정성과 성능 유지, 장애 대응, 보안 관리 등을 포함합니다. 정보 기술 인프라의 안정성과 신뢰성을 유지하기 위해 중요한 역할을 합니다.

클라우드 컴퓨팅
Cloud Computing 인터넷을 통해 서버, 스토리지, 네트워크 등의 컴퓨팅 리소스를 온디맨드로 제공하는 기술로, 기업은 인프라를 보유하거나 유지보수할 필요 없이 필요한 서비스를 유연하게 이용하고 비용을 절감할 수 있습니다. 사용자는 인터넷을 통해 원격으로 액세스하여 필요한 컴퓨팅 리소스를 빠르게 구축하고 관리할 수 있으며, 클라우드 제공업체는 확장성과 가용성을 확보하여 안정적인 서비스를 제공합니다. 클라우드 컴퓨팅은 기업들의 디지털 전환과 혁신적인 애플리케이션 개발에 중요한 역할을 하고 있습니다.

프로그래밍

프로그래밍 언어 컴퓨터는 0과 1(이진수) 인식하여 작동하기 때문에 과거에는 0과 1을 활용한 기계어로 프로그래밍을 했습니다. 하지만 0과 1을 활용한 기계어는 매우 어렵고 복잡했기 때문에 보다 쉬운 방법이 연구되었습니다. 이런 배경 가운데 탄생한 것이 프로그래밍 언어입니다. (C, 자바, 자바스크립트, 파이썬, C++ 등)

파이썬 간결한 문법으로 입문자가 이해하기 쉽고, 다양한 분야에 활용할 수 있기
Python 때문이다. 백엔드 개발과 데이터 분석, AI 개발을 할 수 있습니다.

자바 객체 지향 프로그래밍 언어로, 한 번 작성한 코드가 다양한 플랫폼에서 실행될 수
Java 있도록 설계되었습니다. JVM(Java Virtual Machine)을 사용하여 중간 언어인 바이트 코드로 변환되며, 플랫폼 독립성을 제공합니다. 멀티스레드 지원과 자동 메모리 관리가 가능하며, 풍부한 라이브러리와 API를 제공하여 개발을 용이하게 합니다. Java는 웹 애플리케이션, 모바일 앱, 대규모 시스템 등 다양한 분야에서 널리 사용됩니다.

컴파일러 프로그래밍 언어로 작성된 소스 코드를 기계어로 변환하여 컴퓨터가 실행할 수
Compiler 있는 프로그램으로 만드는 소프트웨어입니다. 컴파일러는 소스 코드를 읽고 분석하여 오류가 있는지 확인한 후, 기계어로 번역하여 실행 파일을 생성합니다. 이렇게 생성된 실행 파일은 반복적으로 실행할 수 있습니다.

디버깅 프로그램이 예상대로 동작하지 않을 때, 오류를 찾고 수정하는 과정을 말합니다.
Debugging 프로그램의 버그나 오류를 식별하고 이를 해결하여 정상적으로 동작하도록 하는 작업으로, 디버깅 과정에는 코드 검사, 로그 분석, 테스트 등이 포함됩니다.

객체 지향 프로그래밍 (Object-Oriented Programming, OOP) 소프트웨어를 개발하는 패러다임 중 하나로, 현실 세계의 객체를 프로그램의 기본 요소로 취급하는 개발 방법론입니다. 클래스와 객체라는 개념을 활용하여 데이터와 해당 데이터를 처리하는 메서드를 함께 묶어서 표현합니다. 객체 지향 프로그래밍은 코드의 재사용성을 높이고 유지보수를 용이하게 하며, 코드의 가독성과 확장성을 개선하는 장점을 가지고 있습니다. 상속, 캡슐화, 다형성 등의 개념을 통해 코드의 모듈화를 강화하고, 큰 규모의 프로젝트를 효과적으로 관리할 수 있습니다. 객체 지향 프로그래밍은 Java, C++, Python 등 다양한 프로그래밍 언어에서 지원되며, 현대 소프트웨어 개발에서 널리 사용되는 중요한 개발 접근 방식입니다.

API (Application Programming Interface) 소프트웨어 간에 상호작용하기 위한 인터페이스를 제공하는 도구나 규칙을 말합니다. 프로그램끼리 데이터를 주고받거나 서비스를 요청하고 제공하기 위해 사용됩니다. API 는 개발자들이 다른 소프트웨어와 연결하여 기능을 확장하거나 서비스를 이용하는데 활용됩니다. 예를 들어, 온라인 지도 서비스에서 지도 데이터를 제공하는 API 를 사용하면 개발자들은 자신의 웹 애플리케이션에 실시간 지도 기능을 추가할 수 있습니다.

API 는 일종의 계약서로, 제공자는 어떤 기능과 데이터를 제공할지 정의하고, 사용자는 그 기능과 데이터를 요청할 수 있도록 인터페이스를 준수해야 합니다. API 는 소프트웨어 간의 효율적인 상호작용을 도와주며, 웹 서비스, 모바일 앱, 클라우드 서비스 등 다양한 응용 분야에서 활발하게 사용됩니다.

자연어처리 (Natural Language Processing, NLP) 인간이 사용하는 언어를 컴퓨터가 이해하고 처리하는 기술로, 텍스트 또는 음성과 같은 자연어 데이터를 분석하고 이해하는 작업을 의미합니다. 문장의 구문 분석, 의미 추론, 감정 분석, 기계 번역 등 다양한 기술을 포함하며, 인간과 컴퓨터 간의 상호작용과 자동화된 자연어 처리 시스템을 개발하는데 활용됩니다. NLP 는 검색 엔진, 가상 비서, 챗봇, 기계 번역, 문서 분류 등 다양한 응용 분야에서 널리 활용되고 있습니다.

웹 개발

프론트엔드 웹 개발은 크게 프론트엔드(Front-end)와 백엔드(Back-end) 기술로 나뉩니다.

Front-end 프론트엔드는 웹페이지 디자인이나 버튼 기능처럼 사용자가 바로 볼 수 있는

백엔드 부분입니다. 백엔드는 사용자가 눈으로 볼 수 없는 뒷단 기술이며 서버단

Back-end 기술이라 불리기도 합니다. HTML, CSS 등이 대표적인 프론트엔드 기술이며, DB 나 서버를 다루는 부분이 백엔드 기술입니다.

UI 컴퓨터나 모바일기계 등을 사용자가 좀더 편리하게 사용할 수 있는 환경을

User Interface 제공하는 설계 또는 그 결과물을 말합니다.

Windows 와 같은 아이콘과 마우스를 통한 작업이 가능한 컴퓨터 인터페이스를 '그래픽 유저 인터페이스(graphic user interface, GUI)'라고 하며 텍스트 입력으로 사용하는 인터페이스를 '캐릭터 유저 인터페이스(Character User Interface, CUI)'라고 합니다.

UX 사용자가 어떤 시스템, 제품, 서비스를 직·간접적으로 이용하면서 느끼고 생각하게

User Experience 되는 지각과 반응, 행동 등 총체적 경험을 말합니다.

프레임워크 소프트웨어 개발을 위한 기반이 되는 구조와 라이브러리의 모음으로, 개발자들이 특정한 문제를 해결하기 위해 사용할 수 있는 미리 정의된 기능들을 제공합니다. 프레임워크는 개발 시간을 단축하고 효율성을 높여주며, 일반적인 작업들을 자동화하고 일관성을 유지할 수 있도록 도와줍니다. 또한, 프레임워크를 사용하면 보다 안정적인 소프트웨어를 개발할 수 있으며, 특정 플랫폼에 종속되지 않는 이식성을 제공합니다.

HTML 모든 웹페이지의 중추입니다. 웹 페이지의 구조를 생성하기 위해 HTML 코드를 작성하는 방법을 알아야 합니다. 여기에는 웹 페이지의 제목, 단락, 목록, 표, 양식 및 기타 필수 요소를 만드는 작업이 포함됩니다.

CSS 웹 페이지의 요소에 스타일을 지정하고 레이아웃을 지정하는 데 사용됩니다. CSS 를 사용하여 웹 페이지의 글꼴 크기, 색상, 배경, 바깥여백, 안쪽여백 및 기타 스타일 속성을 변경하는 방법을 알아야 합니다. 또한 CSS 를 사용하면 다양한 화면 크기와 장치에 적응할 수 있도록 웹 페이지를 반응형으로 만들 수 있습니다.

JavaScript

웹 페이지에서 사용자로부터 특정 이벤트나 입력 값을 받아 동적인 처리를 목적으로 고안된 객체 기반의 스크립트 프로그래밍 언어입니다. 사용자 경험을 향상시키기 위한 방법으로 주로 사용된다. 일반적으로 HTML 문서에 내재되며, 브라우저에서 실행됩니다.

예를 들면, 마우스를 HTML 문서의 특정 이미지나 텍스트 위로 가져갔을 때 이미지 크기가 커지거나 텍스트의 색이 바뀌는 기능이다. 또 검색창 입력 박스에 '자'를 입력하면 현재 기준으로 사용자 검색이 많은 '자'로 시작하는 검색어 목록을 보여주는 기능도 자바스크립트로 구현한 것입니다. 최근에는 이 자바스크립트 언어로 쉽게 프론트엔드와 백엔드 개발을 할 수 있도록 도와주는 프레임워크들이 등장하여 인기 수직 상승 중인 언어입니다.

JSP

JavaServer Pages

서버 측에서 동적 웹 페이지를 생성하기 위한 웹 프로그래밍 기술로, HTML 내에 Java 코드를 삽입하여 동적으로 콘텐츠를 생성할 수 있도록 해줍니다. JSP는 Java 언어를 기반으로 하며, Java Servlet과 함께 사용되어 웹 애플리케이션을 개발하는데 주로 활용됩니다. JSP를 사용하면 서버에서 데이터를 가공하고 웹 페이지를 생성하여 클라이언트에게 동적인 콘텐츠를 제공할 수 있습니다.

React

단일 웹 페이지나 모바일 앱에서 사용자 인터페이스 중 화면에 표시되는 뷰 부분의 개발에 사용되는 자바스크립트 라이브러리입니다.

페이스북(Facebook)에서 개발하여 2013년 처음 배포한 오픈소스입니다. 모델 · 뷰 · 컨트롤러(MVC: Model-View-Controller) 개발 요소 중 뷰(view)만 제공하며, 기존의 웹 기술 HTML, CSS 등과 결합하여 사용할 수 있어 확장성이 뛰어납니다. 개발자는 자바스크립트만으로도 개발할 수 있어 사용이 편리하다.

Node.js

자바스크립트는 프론트엔드 분야에서 주로 활용됐습니다. 그런데 '노드 JS(Node JS)'가 등장하면서 그 용도가 전혀 달라졌습니다. 노드 JS라는 프레임워크 덕분에 자바스크립트로 서버단 기술까지 제어할 수 있게 된 것입니다. 2009년 혜성같이 등장한 노드 JS는 웹 개발자에게 많은 관심을 받고 있으며, 최근 재단까지 설립되며 기술 개발이 활발히 이뤄지고 있습니다.

Django

파이썬 기반의 오픈 소스 웹 애플리케이션 프레임워크로, 웹 사이트 개발을 간편하게 만들어주는 도구입니다. 데이터베이스 모델링, URL 라우팅, 템플릿 엔진, 보안 기능 등 다양한 기능을 내장하고 있어, 빠르고 안정적인 웹 애플리케이션을 구축할 수 있습니다.

Java Spring

Java Spring은 Java 기반의 오픈 소스 애플리케이션 프레임워크로, 기업급 웹 애플리케이션 개발에 사용됩니다. 의존성 주입, AOP(Asspect-Oriented Programming), 보안 기능 등을 제공하여 개발자들이 효율적으로 웹 애플리케이션을 구축할 수 있도록 돕습니다.

Vue.js Vue.js 는 자바스크립트 기반의 프론트엔드 프레임워크로, 사용자 인터페이스를 만들기 위해 사용됩니다. 가볍고 빠르며 컴포넌트 기반 구조를 갖추어 유지보수가 용이하며, 반응형 웹 애플리케이션을 개발하는데 적합합니다.

TypeScript TypeScript 는 자바스크립트의 확장된 언어로, 정적 타입 시스템을 도입하여 더 강력한 타입 체크를 제공합니다. 자바스크립트 코드를 호환하면서도 개발자들에게 코드 작성 시 타입 오류를 미리 발견하여 코드 품질을 향상시키는데 도움을 줍니다.

MongoDB MongoDB 는 NoSQL 데이터베이스로, 문서 지향적인 접근 방식을 사용하여 유연하고 확장 가능한 데이터 저장 및 관리를 제공합니다. JSON 과 비슷한 BSON 형식으로 데이터를 저장하며, 다양한 플랫폼에서 사용 가능하며, 대용량 및 실시간 데이터 처리에 특화되어 있습니다. SQL 기반의 관계형 데이터베이스와는 다른 데이터 구조와 질의 언어를 사용하며, 클라우드 기술과도 통합이 가능하여 유연하고 확장 가능한 애플리케이션을 개발하는 데 적합합니다.

렌더링 렌더링은 웹 페이지 또는 애플리케이션의 내용을 웹 브라우저에 표시하는 과정을 의미합니다. 서버 측 렌더링과 클라이언트 측 렌더링이 있으며, 서버 측 렌더링은 서버에서 페이지를 구성하고 클라이언트에게 보내는 방식이고, 클라이언트 측 렌더링은 웹 브라우저에서 자바스크립트를 사용하여 동적으로 페이지를 생성하는 방식입니다.

컴포넌트 Component 웹페이지의 독립적인 구성 요소로, 재사용 가능한 기능과 정보를 가지며, 주로 HTML, CSS, JavaScript 로 만들어집니다. 이 컴포넌트들은 웹페이지의 다양한 부분에 적용되어 일관성 있고 효율적인 웹 애플리케이션을 구축하는데 사용됩니다. 컴포넌트 기반 웹 개발은 유지보수와 확장성을 향상시키는데 도움이 되며, 프론트엔드 프레임워크를 활용하여 개발하는 것이 보편적입니다.

Servlet 자바를 기반으로 웹 애플리케이션을 개발하기 위한 기술로, 웹 서버에서 동작하는 자바 클래스를 의미합니다. 클라이언트로부터 HTTP 요청을 받아들이고, 그에 따른 동적인 응답을 생성하는 역할을 수행합니다. 웹 애플리케이션 서버에서 동작하며, 자바 서블릿 컨테이너에서 실행됩니다. 웹 애플리케이션 개발 시, 서블릿을 활용하여 사용자 요청을 처리하고 데이터베이스와 상호작용하여 동적인 웹 페이지를 생성하는 등 다양한 기능을 구현할 수 있습니다.

Weka Weka 는 무료 오픈소스 기계 학습 및 데이터 마이닝 도구로, 데이터 분석, 모델 생성, 시각화 등 다양한 작업을 지원하며 사용자 친화적인 그래픽 인터페이스를 제공합니다. Java 기반으로 구현되어 플랫폼에 독립적이며, 데이터 형식 다양성과 데이터 시각화 기능으로 유용하게 활용됩니다. 학계와 산업 현장에서 널리 사용되며, 데이터 과학자와 분석가들에게 다양한 문제 해결에 도움을 줍니다.

데이터 분석·시각화

빅데이터
Big Data 대량의 다양한 종류의 데이터로부터 가치를 추출하는 기술과 방법론을 의미합니다. 이러한 데이터는 정형, 비정형, 반정형 데이터로 구성되며, 기존 데이터베이스 관리 도구로 처리하기 어려운 규모와 복잡성을 가지고 있습니다. 빅데이터를 분석하고 활용함으로써 새로운 인사이트와 패턴을 발견하며, 예측과 의사결정에 유용한 정보를 제공합니다. 이를 통해 기업과 조직은 비즈니스 성과를 향상시키고 혁신적인 서비스를 개발하는 데에 활용하고 있습니다.

EDA
(Exploratory Data Analysis) 데이터 분석 첫 단계에서 많이 진행하는 작업입니다. 주어진 데이터 셋의 특성을 분석하고 간단한 시각화를 통해서 인사이트를 뽑아내는 단계입니다.

전처리 preprocessing 이라고도 부릅니다. 원본 데이터는 누락되거나 잘못된 값들이 포함되어 있을 수 있습니다. 이를 그대로 사용해서 예측 모델을 만들 경우, 결과가 잘 안나올 수 있습니다. 때문에 미리 데이터들을 한번 정제하는 작업을 거치는데 이를 전처리라고 부릅니다.

시각화
Visualization 데이터의 분포나 추이 변화들을 직관적으로 보여주기 위해 시각화하는 것을 말합니다. 막대 그래프, 원 그래프, 스캐터 플랏, 박스 플랏, 히스토그램, 라인 그래프 등이 있습니다.

머신러닝
Machine Learning 인공 지능의 한 분야로, 컴퓨터 시스템에 데이터를 기반으로 학습하고 패턴을 발견하여 문제를 해결하거나 예측하는 기술을 의미합니다. 기계가 일정한 규칙을 프로그래밍하지 않고 데이터로부터 스스로 학습하고 지식을 습득하여 인간과 유사한 의사결정을 수행할 수 있습니다. 머신 러닝은 이미지 인식, 자연어 처리, 추천 시스템 등 다양한 분야에서 활용되며, 데이터의 양과 복잡성이 증가하는 현대 사회에서 중요한 기술로 자리잡고 있습니다.

데이터마이닝
Data Mining 대규모의 데이터에서 유용한 정보와 패턴을 추출하는 기술로, 통계, 머신 러닝, 인공 지능 등을 활용하여 데이터의 숨겨진 지식을 발견하는 과정을 의미합니다. 데이터 마이닝은 데이터베이스, 데이터 웨어하우스, 빅데이터 등 다양한 소스로부터 데이터를 수집하고 전처리한 후, 클러스터링, 분류, 예측, 연관 규칙 발견 등의 기술을 적용하여 의사결정을 지원하고 비즈니스 인텔리전스를 제공합니다. 이를 통해 기업은 고객 행동을 분석, 제품 개발에 필요한 인사이트를 얻거나 사기 탐지와 같은 다양한 분야에서 효과적인 솔루션을 구축할 수 있습니다.

데이터 분석·시각화 - 데이터베이스 관련

관계형 데이터베이스 관리 시스템 RDBMS	SQL 데이터베이스라고도 부릅니다. 데이터들을 테이블 형태의 표로 저장하고, 테이블 간에 서로 관계를 설정할 수 있는 데이터 베이스입니다. SQL 언어를 사용해서 데이터를 저장, 가공, 추출 등을 할 수 있습니다. 가장 범용적으로 많이 사용되며, 데이터를 안전하게 저장하고, 다양한 기능들을 구현할 수 있습니다.
NoSQL (Not Only SQL)	RDBMS의 한계를 극복하기 위해 등장한 데이터 베이스들입니다. 테이블로 표현하기 복잡한 데이터들을 쉽게 저장할 수 있게 지원한다던가, 아주 단순한 형태로 데이터를 저장해서 성능을 극대화하는 경우도 있습니다. 현업에서는 RDBMS를 기본 데이터 베이스로 사용하고, 성능 저하가 심한 기능이나 데이터를 NoSQL를 활용하여 처리합니다.
SQL	Structured Query Language의 약자로 RDBMS에 데이터를 저장하고 읽어오기 위해 반드시 사용해야하는 질의 언어입니다.
스키마 Schema	데이터베이스에서 저장되는 데이터의 구조와 제약 조건을 정의하는 청사진으로, 테이블의 열명, 데이터 타입, 관계 등을 결정합니다. 데이터 일관성과 무결성을 유지하며, 데이터베이스 객체의 정의와 보안 설정을 포함하며, 데이터베이스 설계를 문서화하여 명확하고 이해하기 쉽게 합니다. 스키마는 데이터베이스 관리 시스템(DBMS)에 의해 관리되며, 논리적 스키마와 물리적 스키마로 나뉘어 데이터 저장 방법과 사용자 인터페이스를 정의합니다.
데이터 모델링 Data Modeling	현실 세계의 데이터를 데이터베이스에 효율적으로 표현하는 과정으로, 데이터의 구조와 관계를 정의하여 데이터베이스 테이블로 나타내는 작업을 말합니다. 이를 통해 데이터의 중복을 최소화하고 일관성을 유지하여 데이터베이스의 성능과 관리를 향상시킬 수 있습니다. 데이터 모델링은 데이터베이스 시스템을 설계하는데 필수적이며, 정확하고 체계적인 데이터 모델링은 효과적인 데이터베이스 시스템을 구축하는데 큰 도움이 됩니다.
Pandas	파이썬 데이터 분석 라이브러리로, 데이터 구조와 처리 기능을 제공하여 테이블 형태의 데이터를 쉽게 다룰 수 있습니다. 데이터프레임을 통해 간편한 데이터 조작과 필터링, 집계, 결측치 처리 등을 제공하며, 데이터 시각화와 데이터베이스 연동 등 다양한 기능을 제공합니다.

데이터 분석·시각화 - 머신러닝 관련

딥러닝
Deep Learning 인공 신경망을 기반으로 한 머신 러닝의 한 분야로, 다층 구조의 신경망을 사용하여 데이터에서 복잡한 패턴과 특징을 자동으로 학습하는 기술입니다. 많은 데이터와 컴퓨팅 자원이 필요하지만, 이미지, 음성, 텍스트 등 다양한 데이터 유형에서 뛰어난 성능을 보여줍니다. 딥러닝은 이미지 인식, 음성 인식, 자연어 처리, 게임과 같은 다양한 분야에서 활용되며, 인간의 뇌의 구조를 모방하여 학습하고 예측하는 특징을 갖고 있습니다.

지도학습
Supervised Learning 학습하고자 하는 데이터에 라벨이 붙어있는 경우를 말합니다. 네이버 영화 리뷰 데이터셋에서 긍정적인 리뷰와 부정적인 리뷰를 분류하는 모델을 학습시킨다고 가정하겠습니다.

Plain Text

CopyCaption

“인생 최고의 영화였습니다.” → 긍정

“내내 졸다가 중간에 나왔습니다.” → 부정

이 처럼 데이터마다 라벨이 붙어있고, 이 라벨에 따라서 AI 모델을 학습시키는 것을 supervised learning 이라고 부릅니다.

비지도 학습
Unsupervised Learning 학습하고자 하는 데이터에 라벨이 없이 학습하는 방식을 말합니다. 요새 화제인 GPT 도 기본적으로 이 unsupervised learning 방식으로 학습합니다.

Plain Text

CopyCaption

인생 ?

인생 최고의 ?

인생 최고의 영화 ?

인생 최고의 영화였습니다 ?

위 예시처럼 라벨이 없이, 앞에 오는 단어들만 보고 뒤에 오는 단어를 예측하는 방식으로 학습시킨 AI 모델이 GPT 입니다. 비지도 학습의 대표적인 예시로는 클러스터링, GAN, difusion 모델 등이 있습니다.

Classification 입력 데이터를 미리 정해진 카테고리 가운데 하나로 분류하는 것입니다.

분류 예를들어 손글씨로 쓰인 숫자 이미지를 입력으로 받았을 때, 무슨 숫자인지 0 부터 9 사이 하나로 분류하는 것이 있습니다.

Regression 입력 데이터를 가지고 숫자로 된 예측 값을 내리는 것을 말합니다. 예를들어

회귀 아이의 출생 연도, 성별, 인종, 부모의 키 등을 입력받아서 예상 신장을 예측하는 것이 있습니다.

Clustering 비슷한 특성의 데이터들끼리 그룹으로 묶어주는 테스트를 말합니다. 예를 들어

클러스터링 얼굴 이미지 데이터 셋이 있다고 했을 때, 클러스터링을 하게 되면 비슷한 생김새의 이미지들끼리 같은 그룹으로 묶이게 됩니다.

Redis 빠르고 간단한 인메모리 데이터베이스로, 데이터를 메모리에 저장해 빠르게 접근할 수 있습니다. 주로 캐싱이나 세션 관리에 사용되며, 문자열이나 리스트와 같은 다양한 데이터를 저장할 수 있습니다. 또한, 데이터의 지속성을 보장하기 위해 디스크에 저장하는 옵션도 제공합니다. 성능이 우수하고 사용이 편리하여 다양한 애플리케이션에서 인기가 있습니다.

NumPy 파이썬의 과학 및 수치 계산을 위한 핵심 라이브러리로, 다차원 배열과 행렬을 다루는 기능을 제공합니다. NumPy 는 빠르고 효율적인 수치 계산을 위해 C 언어로 구현되었으며, 다양한 수학 함수와 브로드캐스팅 기능을 지원하여 배열 데이터를 효과적으로 다룰 수 있습니다. 이러한 기능들은 데이터 분석, 머신 러닝, 이미지 처리 등 다양한 과학적이고 엔지니어링 분야에서 많이 활용됩니다. NumPy 는 파이썬의 기본 데이터 구조보다 빠르고 메모리 효율적이기 때문에 대용량 데이터 처리에 유용하며, 많은 데이터 분석 라이브러리들이 NumPy 를 기반으로 구축되어 있습니다.

데이터 분석·시각화 - 딥러닝 관련

CNN (Convolution Neural Network)	인간의 시신경이 동작하는 방식과 유사하게 동작하는 신경망 기본 블록입니다. 이를 층층이 쌓아서 다양한 네트워크를 구성하여 이미지 분류, 물체 인식, 얼굴 인식 등에 사용합니다.
RNN (Recurrent Neural Network)	인간의 언어처럼 데이터의 등장 순서가 중요한 데이터들이 있습니다. 이를 sequential data 라고 표현합니다. 이 sequential data 를 처리하기 위해서는 이전까지의 맥락을 기억하는 것이 중요한데, 이를 지원하는 신경망 구조가 RNN 입니다.
LSTM	RNN 의 일종으로 시퀀스가 길어지더라도 맥락을 잘 기억할 수 있도록 개량한 딥러닝 기본 블록입니다. 기계 번역, 문장 생성 등에 사용됩니다.
GRU	LSTM 과 비슷하지만 연산을 더 단순화 시킨 블록입니다. LSTM 과 마찬가지로 기계 번역, 문장 생성 등에 사용합니다.
Transformer	수업을 통해 배우겠지만, RNN 은 근본적으로 학습시킬 때 비효율이 발생합니다. 때문에 대량의 텍스트를 학습하는 것에 한계가 있습니다. 이러한 문제점을 극복하고, 맥락을 효과적으로 기억할 수 있는 self-attention 이라는 메커니즘을 중심으로 만든 신경망 구조가 transformer 입니다. 최근에는 컴퓨터 비전, 자연어 처리, 음성 합성 등 거의 모든 분야에서 이 트랜스포머를 기반으로 한 AI 모델들이 뛰어난 성능을 보여줍니다.

그 외 주요 서비스

Git 컴퓨터 파일의 변경사항을 추적하고 여러 명의 사용자들 간에 해당 파일들의 작업을 조율하기 위한 스냅샷 스트림 기반의 분산 버전 관리 시스템입니다.

GitHub 컴퓨터 프로그램 소스를 공유하고 협업하여 개발할 수 있는 버전 관리 시스템 깃(Git)에 프로젝트 관리 지원 기능을 확장하여 제공하는 웹 호스팅 서비스입니다.

Notion 노션은 대화식 작업 및 협업 톨로, 개인 및 조직에서 사용할 수 있습니다. 개인,팀,그리고 조직이 지식을 공유하고, 업무를 협업하며, 생산성을 높이는 데에 도움이 됩니다. 노션은 무료 버전과 유료 버전이 있으며, 무료 버전에도 다양한 기능을 사용할 수 있습니다.

ChatGPT . OpenAI 에서 개발한 대화형 인공지능 모델로, 사용자와 자연스럽게 대화하며 문맥을 이해하여 응답을 생성하는 기술입니다. GPT 모델의 기반인 트랜스포머(Transformer) 구조를 사용하여 대화 데이터를 학습하고, 다양한 자연어 처리 작업에 뛰어난 성능을 보여줍니다. ChatGPT 는 대화형 서비스와 챗봇 등 다양한 분야에서 사용되며, 자연어 처리 기술의 발전에 큰 영향을 미치고 있습니다.

JIRA 소프트웨어 개발과 프로젝트 관리에 사용되는 강력한 협업 도구입니다. 이슈 트래킹 시스템으로, 팀원들이 작업 항목을 등록하고 추적하여 프로젝트의 진행 상태를 파악할 수 있습니다. 팀원들은 할당된 작업에 대한 업무를 확인하고 업데이트하며, 각 작업은 이슈라는 티켓으로 표시됩니다. Jira 는 대시보드와 보고서를 제공하여 프로젝트 팀이 현재 상태와 진행 사항을 시각적으로 파악할 수 있도록 도와주며, 스프린트 기능을 통해 애자일 방식의 프로젝트 관리를 지원합니다. 또한, 다양한 플러그인과 통합 기능을 제공하여 다른 도구들과의 연동이 용이하며, 개발자들의 협업과 업무 관리를 효율적으로 지원하는 강력한 플랫폼입니다.

AWS 아마존이 제공하는 클라우드 컴퓨팅 플랫폼으로, 인프라스트럭처 서비스, 플랫폼 서비스, 소프트웨어 서비스 등 다양한 클라우드 기반 서비스를 제공하여 기업들이 유연하고 확장 가능한 IT 인프라를 구축하고 운영할 수 있도록 합니다. AWS 는 빠르고 안정적인 클라우드 인프라를 제공하여 애플리케이션을 빠르게 개발, 배포하고 비용을 절감할 수 있으며, 인공지능, 빅데이터, 머신 러닝 등의 기술과 서비스를 통합하여 혁신적인 솔루션을 구현할 수 있습니다.

Azure 마이크로소프트가 제공하는 클라우드 컴퓨팅 플랫폼으로, 컴퓨팅, 스토리지, 빅데이터, 인공지능, 머신 러닝 등 다양한 서비스를 포괄하는 종합적인 클라우드 솔루션을 제공합니다. 기업들은 Azure 를 활용하여 애플리케이션을 더 빠르고 안정적으로 개발, 배포하고 필요한 리소스를 탄력적으로 조정할 수 있으며, 보안성과 컴플라이언스를 강화하여 신뢰성 높은 서비스를 제공합니다. 또한, 개방적인 생태계와 통합 기능을 제공하여 다양한 데이터 소스와 플랫폼 간의 연결성을 강화하고 혁신적인 디지털 솔루션을 구현할 수 있습니다.

GCP 구글이 제공하는 클라우드 컴퓨팅 플랫폼으로, 컴퓨팅, 스토리지, 빅데이터, 인공지능, 머신 러닝 등 다양한 클라우드 서비스를 제공합니다. 기업들은 GCP 를 활용하여 확장 가능하고 안정적인 클라우드 인프라를 구축하며, 웹 애플리케이션, 모바일 앱, 데이터 분석 등의 혁신적인 서비스를 개발할 수 있습니다. GCP 는 오픈 소스 기반의 플랫폼과 다양한 통합 기능을 제공하여 개발자들의 창의성을 이끌어내고, 다양한 비즈니스 요구를 충족시키며 혁신적인 디지털 솔루션을 구현할 수 있습니다.

SaaS (Software as a Service) 클라우드 컴퓨팅에서 제공되는 소프트웨어 애플리케이션으로, 웹 브라우저를 통해 인터넷을 통해 접근하고 사용할 수 있습니다. 사용자는 서버나 인프라를 구축할 필요 없이 웹에서 애플리케이션을 이용할 수 있으며, 업데이트와 유지보수는 제공업체가 담당합니다.

PaaS (Platform as a Service) 클라우드 컴퓨팅에서 제공되는 애플리케이션 개발 및 배포를 위한 플랫폼으로, 인프라 구축에 신경 쓸 필요 없이 애플리케이션을 개발하고 실행하는 환경을 제공합니다. 개발자들은 애플리케이션 코드에 집중할 수 있으며, 플랫폼은 자동으로 스케일링, 로드 밸런싱 등을 관리합니다.

IaaS (Infrastructure as a Service) 클라우드 컴퓨팅에서 제공되는 가상화된 컴퓨팅 리소스, 스토리지, 네트워크 등의 기본 인프라를 말합니다. 사용자는 가상 머신, 스토리지, 네트워크를 필요에 따라 선택하여 사용하며, 인프라에 대한 관리와 운영 책임은 사용자가 가지고 있습니다. IaaS 는 유연하고 확장 가능한 인프라를 제공하여 다양한 애플리케이션을 지원합니다.