

## 1페이지

-제목 : 기계학습입문

-약력 : 배상민 강사

한국과학기술원(KAIST) 산업 및 시스템 공학과 학/석사 전공

한국과학기술원 김재철 AI대학원 박사과정

kakao 추천팀 인턴 경력

Samsung, LG 대기업 강사 경력

## 2페이지

강의 소주제 3가지

1. 기계학습의 첫걸음, 기초 개념 강좌
2. 전문가 양성을 위한, 알고리즘의 심화 이론까지
3. 기계학습의 응용, 실전 프로그래밍 강좌

## 4페이지

강의 주 시청자

1. 대학생/구직자

기계 학습에 대한 기초부터 심화 이론까지 이해함으로써, 최근 기업들이 원하는 인공 지능 전문가로서의 역량을 키울 수 있습니다.

2. 직장인

기계 학습의 원리에 대한 깊은 이해를 바탕으로, 현업에 필요한 기계 학습 모델을 설계 및 구현할 수 있습니다.

## 5페이지

강의에서 활용하는 툴과 알고리즘

핵심 툴 : 파이썬

1. 파이썬
2. 넘파이 (numpy)
3. 판다스 (pandas)
4. 사이킷런 (sklearn)

## 8페이지

이론 예제 그림 폴더에 넣어두었습니다.

## 9페이지

강의를 통해 얻을 수 있는 4가지 항목 20~30자 분량

1. 다양한 알고리즘 개념이해  
기계 학습의 다양한 알고리즘들에 대한 기초 원리 이해
2. 심화 이론 이해  
AI 전문가에게 필요한 수학적 심화 이론까지 습득
3. 기계 학습 모델 설계

해결하고 싶은 문제에 대한 최적의 기계 학습 모델을 제안/설계

#### 4. 실전 프로그래밍

실제 데이터를 활용한 다양한 알고리즘 구현

10페이지

강의 커리큘럼 챕터별 요약 3개 내외

#### 1. 회귀 모델

이론 :

- 머신 러닝의 정의, 경사 하강법, 편향과 분산 오류

실습 :

- colab 연동, numpy와 sklearn을 활용한 경사 하강법 구현, 선형 회귀 모델 학습

과제 :

- numpy와 sklearn을 통한 다항 회귀 모델 구현(과제)

#### 2. 분류 모델

이론 :

- 로지스틱 회귀, 소프트맥스 회귀, 서포트벡터 머신, 결정 트리, 선형 판별 분석

실습 :

- 로지스틱 회귀, SVM, Decision Tree 모델 구현 및 학습, 학습된 모델과 데이터에 대한 시각화

과제 :

- 하이퍼파라미터 서치 툴을 이용한 svm 학습 및 분석(**프로젝트**)
- 최적의 결정 트리 모델 학습 (과제)

#### 3. 앙상블 학습

이론 :

- 배깅, 부스팅

실습 :

- 랜덤 포레스트, AdaBoost 구현

과제 :

- Gradient Boosting, XGBoost, Light GBM 구현 (과제)
- House Price 캐글 문제 해결 (**프로젝트**)

#### 4. 차원 축소

이론 :

- PCA, SVD, LDA, t-SNE, UMAP

실습 :

- 비지도 학습의 차원 축소(PCA, t-SNE), 라벨을 활용한 차원 축소(LDA), 데이터 압축(SVD)

과제 :

- UMAP을 통한 숫자 데이터셋 학습 모델 구현 (과제)

## 5. 군집화

이론 :

- K-Means, Mean Shift, DBSCAN, GMM, Hierarchical Clustering

실습 :

- 군집화 알고리즘 K-Mean, DBSCAN 구현 및 시각화

과제 :

- MeanShift와 GMM 군집화 알고리즘 구현 및 분석 (**프로젝트**)

## 13페이지

### Chapter 1. 회귀 모델

#### 1.1 머신 러닝 소개

머신 러닝의 개념과 분야들에 대해 알아본다.

#### 1.2 기초 수학

머신 러닝을 이해하기 위해 필요한 기초 수학 지식을 알아본다.

#### 1.3 경사 하강법

회귀 모델을 학습하기 위해 손실 함수를 정의하고, 경사 하강법을 알아본다.

#### 1.4 편향과 분산

편향과 분산 오류 문제를 이해하고, 이를 해결하기 위해 교차 검증, 정규화 손실 함수에 대해 알아본다.

### Chapter 2. 분류 모델

#### 2.1 로지스틱 회귀

로지스틱 회귀 모델과 최대 우도 추정법을 통한 손실 함수를 알아본다.

#### 2.2 소프트맥스 회귀

다중 클래스의 상황에서 사용되는 소프트맥스 함수와 교차 엔트로피에 대해 알아본다.

#### 2.3 서포트벡터 머신

SVM 모델의 최적화 문제를 풀어보고, 커널 함수에 대해 알아본다.

#### 2.4 결정 트리

회귀와 분류에서의 결정 트리를 배워보고, 트리 모델의 장단점에 대해 알아본다.

#### 2.5 선형 판별 분석

베이즈 분류기와 선형 판별 함수에 대해 알아본다.

### Chapter 3. 앙상블 학습

#### 3.1 배깅

부트스트랩의 개념과 배깅 모델, 랜덤 포레스트에 대해 알아본다.

#### 3.2 부스팅

AdaBoost, Gradient Boosting 등의 다양한 알고리즘들의 원리에 대해 알아본다.

### Chapter 4. 차원 축소

#### 4.1 PCA

주성분 분석(PCA)의 개념과 고유값 분해에 대해 알아본다.

#### 4.2 SVD

특이값 분해(SVD)와 데이터 압축에서의 활용을 알아본다.

#### 4.3 선형 판별 분석

이전의 베이지스 분류기가 아닌 분산의 관점에서 차원 축소에서의 LDA에 대해 알아본다.

#### 4.4 t-SNE

t-SNE 알고리즘의 학습 방법과 t 분포에 대해 알아본다.

#### 4.5 UMAP

t-SNE 알고리즘의 한계와 최신 알고리즘 UMAP에 대해 알아본다.

### Chapter 5. 군집화

#### 5.1 K-means

K-means 군집화와 EM 알고리즘에 대해 알아본다.

#### 5.2 Mean Shift

커널 밀도 추정과 Mean Shift 군집화에 대해 알아본다.

#### 5.3 DBSCAN

밀도 기반의 군집화 방식의 DBSCAN과 그의 장점에 대해 알아본다.

#### 5.4 Gaussian Mixture Model

모수적 추정을 통한 GMM 알고리즘의 학습 방법에 대해 알아본다.

#### 5.5 계층적 군집화

데이터 세분화에 사용되는 계층적 군집화 방법에 대해 알아본다.