



바이트디그리
Machine Learning with Python

Syllabus
Fast Campus
2020. 08

00. Byte Degree Description

바이트 디그리란, 직장인 실무 교육을 운영하는 패스트캠퍼스가 산업 전문가들과 협력을 통해 구성한 기술 인증 교육 코스입니다. 디지털 기술의 빠른 변화 속에서 현업 전문가가 설계한 프로젝트를 기반으로 학습하고, 단순히 귀와 눈으로만 익히는 것에서 벗어나 수강생 손으로 직접 구현해보며 해당 기술을 익히고, 프로젝트를 통해 이해한 정도를 검증할 수 있도록 합니다. 수강생이 학습한 내용을 바탕으로 제출하는 프로젝트는 리뷰어에게 1:1 로 피드백을 받으며 일정 수준을 통과해야 바이트 디그리를 받을 수 있습니다. 전통적인 대학 학위보다 적은 시간을 들여 실무에 가장 가까운 기술을 익히고, 프로젝트 결과물을 통해 이해도를 검증하고, 인증받을 수 있습니다.

01. Course Description

Byte Degree : 머신러닝은 [올인원 패키지 : 머신러닝과 데이터분석 A-Z] 에서 꼭 알아야하는 파트, 머신러닝을 집중적으로 다루는 파트들을 추출해 커리큘럼으로 구성하였습니다.

또한 머신러닝을 학습하고자 하는 입문자분들도 기초 확률/통계부터 학습해나가면서 실습에 앞서 충분한 개념 이해가 가능하도록 설계하였습니다.

더 나아가 핵심적으로 사용되는 머신러닝 알고리즘들을 살펴보고 원리와 개념을 탄탄하게 학습하면서 실습과정인 프로젝트를 진행하기에 앞서 충분한 이론 속지가 가능하겠습니다.

본 강의는 Intro - 머신러닝 개념 부터 수강하게 됩니다. 머신러닝이 무엇인지, 지도학습과 비지도학습이 무엇인지를 이해하고 앞으로 배우게 될 머신러닝 학습의 방향에 대해 미리 살펴보도록 합니다.

Part 1. **[머신러닝을 위한 수학]**에서는 본격적으로 머신러닝을 학습하기 이전에 꼭 필요한 확률/ 통계/ 미분과 같은 수학기초 개념을 학습하게 됩니다.

본 파트는 머신러닝에 꼭 필요한 수학들로만 구성하여 다음 파트부터 머신러닝 알고리즘들을 다루는 데에 초석이 될 것입니다.

Part 2. **[머신러닝의 시작 - 회귀분석]** 파트에서는 통계기법이면서 머신러닝의 가장 기본적인 모형인 '회귀분석'을 배워보고 머신러닝의 개념에 대해 이해해봅니다.

실습도 함께 진행하며 그저 통계 '회귀분석'이 아닌 머신러닝을 위한 '회귀분석'을 학습함으로써 지도학습의 원리를 이해하고, 더 나아가 A or B로 나누는 Binary 형태의 출력변수도 다뤄볼 수 있는 로지스틱 회귀분석까지 학습하게 됩니다.

Part 3. **[머신러닝 기본모형]** 파트부터 본격적으로 머신러닝 알고리즘들을 다양하게 다루게 되며 핵심적인 모형들의 원리와 개념을 이해해봅니다.

특히 핵심적인 지도학습 모델들을 다뤄보면서 머신러닝 대회/데이터분석 대회에서 가장 많이 사용되는 분석기법들을 배워볼 수 있도록 합니다.

Part 4. **[앙상블]** 파트에서는 앞에서 배운 모델들을 활용해 더 나은 성능을 도출하는 방식을 배우게 됩니다. 여러 모델들의 평균값을 통해 모집단의 평균에 가깝게 설계하는 앙상블 방식의 유용성을 이해하고, 특성이 얼마나 적합한지 평가할 수 있는 'shape value' 도 알아보도록 합니다.

Part 5. **[군집분석 & 불균형데이터 다루기]** 파트에서는 '비지도학습'을 학습합니다. 앞에서 핵심적인 지도학습 기법들을 다뤄봤다면, 'Clustering'을 통해 비지도학습의 개념과 활용방식을 배워보고 어떤 데이터에 비지도학습을 사용할 수 있는지 이해합니다. 또한 본 파트에서는 우리가 머신러닝을 학습하면서 자주 마주하게 될 '불균형 데이터'도 다뤄볼 예정입니다. 특히 Oversampling/Undersampling/Hybridresampling 세 가지 방법에 대해 배워보고 불균형 데이터에 대해 어떻게 대처할 수 있는지 이해합니다.

Intro 강의를 포함해 Part1 ~ Part 5 까지는 각각의 개념에 대해 이해도를 평가하기 위한 퀴즈를 실시합니다.

각 주차별로 간단히 배운 내용을 복습하는 퀴즈를 진행하고, Part1~5까지는 각 파트가 끝날 때마다 이해도 평가를 위한 응용_미니프로젝트(6회)를 진행하게 됩니다.

강의가 마무리된 이후 2주 간의 최종 프로젝트를 2회(총 4주)진행함으로써 머신러닝 과정에 대한 실무 역량을 확인하게 됩니다.

02. Course Information

1. 모든 과정(강의 수강, 질의 응답, 프로젝트 피드백)은 온라인으로 진행됩니다.
2. 퀴즈 점수(출석 점수), 파트별 미니 프로젝트(6개), 최종 프로젝트 점수(2개)를 종합하여 평가 합니다.
3. 조교를 통한 질의응답과 프로젝트 결과물에 대한 리뷰어의 꼼꼼한 피드백을 진행합니다.

03. Course Objectives

1. 확률, 통계부터 머신러닝에 꼭 필요한 수학들을 학습하게 되면서 기초를 다집니다.
2. 주로 쓰이는 머신러닝 알고리즘들을 다양하게 학습해보면서 머신러닝의 개념을 이해하고 앙상블, 클러스터링 모형을 통해 머신러닝의 대표적인 구분인 지도학습/비지도학습을 학습합니다.
3. 머신러닝을 실제 사용하면서 자주 마주하게 되는 '불균형데이터'들까지 미리 다뤄보면서 문제해결 역량을 길러보고, 실전 프로젝트를 통해 4주간 이를 검증합니다.

04. Course Task

1. 퀴즈 : 매주(12주) 한번씩 퀴즈를 통해 학습 한 내용을 점검합니다. 퀴즈는 객관식, 단답식, O/X 등 다양한 형태로 출제되며 학습한 내용을 복습하는 정도의 수준이 출제됩니다. 10점 만점으로 출제가 되며 5점 이상 득했을 시, 한 주 출석이 인정됩니다. 총 수업 주차(12주) 중 2/3(8주) 출석 인증이 되어야 바이트 디그리 수료 조건을 충족합니다.
2. Mini Project : Intro 파트를 포함해 총 6개의 파트로 구성되어 있는 [바이트 디그리 : 머신러닝(Python)] 는 Intro 파트만 제외하고 1~5 파트에 대한 이해도를 점검하는 수준의 미니 프로젝트가 출제됩니다. 총 출제 되는 Mini Project(6개) 중 4개 이상 제출하여야 바이트 디그리 수료 조건을 충족합니다.
3. Final Project : 마지막 파트의 수업이 끝날 때 출제되는 프로젝트입니다. Mini Project와 동일한 형태로 문제가 출제 되지만, 모든 파트에 대한 종합적 이해가 요구되는 수준으로 Mini Project 보다 난이도 있게 출제됩니다. 제출한 프로젝트에 대하여 두 개의 Final Project 모두 리뷰어가 정해진 점검표에 맞추어 평가를 진행하고, 최종적으로 통과 하였을 시, 바이트 디그리가 수여됩니다.

05. What You Will Learn

1. 꼭 필요한 수학, 통계 개념을 함께 학습하여 머신러닝의 기저 원리를 이해하게 됩니다.
2. Python 프로그래밍 언어를 활용해 머신러닝 알고리즘을 실제로 사용해보고, 객관적인 지표에 기반한 문제해결 방식을 학습합니다.
3. 다양한 머신러닝 프로젝트를 수행해보면서 이론 뿐만 아니라 실제 머신러닝 활용 및 최적화 방법론을 배우게 됩니다.

06. Course Summary

구분	파트	파트 목표	강사	감수자
Intro	머신러닝 개요	<p>머신러닝은 크게 지도학습과 비지도학습으로 나눌 수 있습니다.</p> <p>1.머신러닝의 기본적인 개념을 학습하고 '입력변수(X)와 출력변수(Y)의 관계를 모델링하는 지도학습'과 '출력변수(Y) 없이 입력변수(X)끼리의 관계를 모델링하는 비지도학습'을 이해합니다.</p> <p>2.다음으로, 앞으로 배우게 될 머신러닝 주요 모형들(회귀분석~클러스터링)은 어떤 것들이 있는지 알아보고 기본적인 내용을 숙지하게 됩니다.</p> <p>3. 마지막으로, 이러한 모형들에 대해 적합성을 평가하는 기준과 Overfitting(과적합)을 방지하기 위한 데이터 분할 방법을 이해합니다. 이를 통해 최적화&모델 성능을 어떻게 높일 수 있는 지 머신러닝의 전반적인 예측 실험 과정을 알아보도록 합니다.</p>	이경택	최윤진 (몬데이터터 CEO / Data Analyst)
Part 1	머신러닝을 위한 수학	<p>머신러닝에 필요한 통계적 개념을 문제 상황에 적용해봅니다.</p> <p>1. '유기질 비료 연구' 예시를 통해 범주형 자료를 이해하고 공분산 및 피어슨 상관계수를 확인합니다.</p> <p>2. '제약회사의 신약 개발'예시를 통해 귀무가설과 대립가설, 그리고 유의수준에 대한 내용을 확인합니다.</p>	김강진	이정훈 (올바른 수학교육 연구소 / 연구소장)
Part 2	머신러닝의 시작, 회귀분석	<p>1.앞서 배운 확률/통계/미분~ Matrix미분까지 활용하여 본격적으로 회귀분석을 학습합니다. 회귀분석을 통해 귀무가설/대립가설~ R Square와 같은 통계개념도 숙지하면서 머신러닝을 위한 회귀분석이 어떻게 코드로 구현되는지 알아봅니다. 또한 Multicollinearity(다중공선성)와 같은 문제현상을 알아보고 VIF 진단법을 학습합니다.</p> <p>2.이렇게 단순/ 다중회귀분석을 학습한 이후 binary 형태의 출력변수를 예측할 때 사용하는 기법인 로지스틱 회귀분석까지 배워봅니다.</p>	이경택	김민중 (국가 수리과학 연구소/ 선임연구원)

Part 3	머신러닝 기본모형	<p>Part 2까지 기초통계~확률 개념과 함께 회귀분석을 배웠다면, Part 3부터는 본격적으로 머신러닝 알고리즘들을 학습해봅니다.</p> <p>1. 특정 환경에서 가질 수 있는 패턴이 두 가지 경우일 때 유용하게 사용할 수 있는 Naive Bayes 2. 정규분포와 같이 규칙적인 분포를 띄지 않는 경우, margin(서로 다른 데이터셋을 나누는 범위)를 최대화하는 boundary를 찾는 SVM(Support Vector Machine) 3. 각 변수들의 이분법 기준을 통해 샘플을 분류하고, 분류된 집단의 성질으로 종속변수를 추정하는 의사결정나무 이 내용들을 수학적 개념과 함께 코드로 구현해보는 실습도 해보면서 머신러닝을 더욱 탄탄하게 이해해보도록 합니다.</p>	김강진	이준호 (J.AI Labs CEO/ Deep Learning Researcher)
Part 4	앙상블	<p>하나의 머신러닝 모델을 사용했을 때 과적합과 같은 문제가 자주 발생하는데, 이를 해결할 수 있는 강력하고 유용한 기법이 'Ensemble' 모델입니다.</p> <p>여러가지 모델의 결과를 평균을 내서 모집단의 평균에 가깝게 만들어내는('대수의 법칙') Ensemble 모델에 대해 이해해보고 Tree 기반의 단일모델인 Bagging, Random Forest, Boosting 과 Regression을 사용하는 Stacking을 알아봅니다. 마지막으로 Feature importance를 파악하는 데에 도움을 주는 'shape value'까지 이해해보도록 합니다.</p>	이경택	이경록 (브레인크루 CEO / Data Scientist)
Part 5	군집분석 & 불균형데이터 다루기	<p>지금까지 배운 머신러닝 모델들은 '지도학습'으로 분류할 수 있습니다. Part 6.에서는 '비지도학습'을 대표하는 군집분석(Clustering)을 배웁니다. 군집분석은 유사데이터간 분류하고 특성을 찾아내는 기법으로, 크게 K-means/Hierarchical/DBSCAN Clustering 세 가지가 있습니다. 모두 배워보면서 실습을 통해 확실하게 개념을 다져보도록 합니다.</p> <p>이렇게 지도학습/비지도학습 알고리즘들을 배워보면서 머신러닝을 이해했다면, 더 나아가 머신러닝을 다루면서 자주 겪게 되는 불균형 데이터 문제를 미리 학습해봅니다. 불균형 데이터를 다루는 기법은 크게 Oversampling/Undersampling/Hybrid resampling 이 있습니다. 세 가지 기법을 각각 알아보고 학습을 마무리합니다.</p>	이경택	김태영 (AIFactory CEO, 캐글코리아&캐라 스코리아 운영진 / Deep Learning Engineer)

07. Course Schedule

- 총 수업 주차 : 16주 (수업 12주, 최종 프로젝트 4주)
- 총 수업 시간 : 1275 m

주차	기간	수업 시간(m)	특이사항
1	2020. 8. 17 ~ 2020. 8. 23	74	Quiz
2	2020. 8. 24 ~ 2020. 8. 30	101	Quiz
3	2020. 8. 31 ~ 2020. 9. 06	113	Quiz / mini Project
4	2020. 9. 07 ~ 2020. 9. 13	171	Quiz
5	2020. 9. 14 ~ 2020. 9. 20	177	Quiz / mini Project
6	2020. 9. 21 ~ 2020. 9. 27	103	Quiz
7	2020. 9. 28 ~ 2020. 10. 04	69	Quiz
8	2020. 10. 05 ~ 2020. 10. 11	92	Quiz / mini Project
9	2020. 10. 12 ~ 2020. 10. 18	150	Quiz
10	2020. 10. 19 ~ 2020. 10. 25	86	Quiz / mini Project
11	2020. 10. 26 ~ 2020. 11. 01	83	Quiz
12	2020. 11. 02 ~ 2020. 11. 08	128	Quiz / mini Project I / mini Project II
최종 프로젝트(A)	2020. 11. 09 ~ 2020. 11. 22	-	Final Project A 제출
최종 프로젝트(B)	2020. 11. 23 ~ 2020. 12. 06	-	Final Project B 제출

08. Course Project

- 총 프로젝트 : 8개 (Mini Project 6개 / Final Project 2개)
- 바이트 디그리 수료 기준 : 6개 (Mini Project 4개 제출 / Final Project 2개 통과)

1. Mini Project

a. 머신러닝을 위한 수학

- 출제자 : 이정훈 (올바른 수학교육 연구소 연구소장)
- 기한 : 2020. 08. 31 - 2020. 09. 14 (2주)
- 프로젝트 명 : 유기질 비료 연구와 제약 실험 사례를 통한 통계적 개념 이해 Project
- 프로젝트 설명
강의에서 배운 머신러닝에 꼭 필요한 통계적 개념을 문제 상황에 적용해봅니다.
먼저 유기질 비료 연구 예시를 통해 범주형 자료를 이해하고 공분산 및 피어슨 상관계수를 확인하게 됩니다.
또한 제약회사의 신약 개발 예시를 통해 귀무가설과 대립가설, 유의수준에 대한 내용 이해를 확인합니다.

b. 머신러닝의 시작, 회귀분석

- 출제자 : 김민중 (국가수리과학연구소 선임연구원)
- 기한 : 2020. 09. 14 - 2020. 09. 28 (2주)
- 프로젝트 명 : MLB 데이터를 활용한 회귀모델 생성 및 결과분석 Project
- 프로젝트 설명 :
이번 파트에서 공부한 회귀분석을 MLB moneyball 데이터에 적용해보는 프로젝트입니다.
단순선형회귀분석과 여러 개의 독립변수를 사용하는 다중선형회귀분석의 결과를 비교할 예정입니다. 또한, VIF를 계산해서 다중공선성 문제를 확인할 것이며 종속변수를 잘 설명하는 최적의 독립변수를 찾는 것을 문제화하였습니다.

c. 머신러닝 기본모형

- 출제자 : 이준호 (/ Deep Learning Researcher)
- 기한 : 2020. 10. 05 - 2020. 10. 19 (2주)
- 프로젝트 명 : 가상데이터를 활용한 분류모델 생성 및 시각화 Project
- 프로젝트 설명 :
머신러닝의 지도학습에는 크게 회귀 문제, 분류 문제가 있습니다. 이번 프로젝트에서는 분류 문제를 해결하는 대표적인 알고리즘 4가지(Naive Bayes, KNN, SVM, 의사결정나무)를 각각 구현해보고 시각화해봅니다.

이를 통해 이번 파트에서 학습한 내용을 토대로 더미(가상)데이터를 직접 다뤄보면서 결과를 예측하고 해석해볼 수 있습니다.

d. 앙상블

- i. 출제자 : 이경록 (브레인크루 CEO / Data Scientist)
- ii. 기한 : 2020. 10. 19 - 2020. 11. 02
- iii. 프로젝트 명 : 앙상블 기법을 활용한 보스턴 주택 가격 예측 (Boston House Prices) Project
- iv. 프로젝트 설명 :
앙상블 파트에서 배운 내용을 통해 보스턴 주택 가격과 관련된 데이터를 기반으로 주택 가격을 예측하게 됩니다. pandas, numpy, scikit-learn 패키지를 활용하여 데이터 가공, 전처리, 피쳐 엔지니어링을 진행하고 회귀문제를 다루는 머신러닝 앙상블 기법을 활용하여 보스턴 주택 가격을 예측합니다. 또한 Bagging, Boosting, Voting, Stacking 기법을 활용하여 앙상블 모델을 구현하고 구현된 각 모델의 성능을 비교해봅니다.

e. 군집분석 & 불균형데이터 다루기 - 군집분석

- i. 출제자 : 김태영 (AI Factory CEO, 캐글코리아&케라스코리아 운영진 / Deep Learning Engineer)
- ii. 기한 : 2020. 11. 02 - 2020. 11. 23
- iii. 프로젝트 명 : 숫자손글씨(MNIST) 데이터 클러스터링 Project
- iv. 프로젝트 설명 :
클러스터링 파트에서는 수치 및 텍스트에 대한 여러가지 클러스터링 기법에 대해서 배워봤습니다. 이번 프로젝트에서는 이미지에 대해서 배운 클러스터링 기법을 활용해보도록 하겠습니다. 데이터는 인공지능 입문에서 가장 많이 사용되고 있는 숫자손글씨인 MNIST에 대해서 다뤄보도록 하겠습니다.

f. 군집분석 & 불균형데이터 다루기 - 불균형데이터 다루기

- i. 출제자 : 김태영 (AI Factory CEO, 캐글코리아&케라스코리아 운영진 / Deep Learning Engineer)
- ii. 기한 : 2020. 11. 02 - 2020. 11. 23
- iii. 프로젝트 명 : IEEE-CIS Fraud Detection의 불균형 데이터 다루기 Project
- iv. 프로젝트 설명 :
실제 캐글에서 개최된 부정거래 예측 대회인 IEEE-CIS Fraud Detection에서 사용된 데이터를 가지고 불균형 데이터 분석을 수행해보고, 간단한 예측 모델을 만들어보겠습니다. 예측 모델의 정확도를 높이기 위해서 본 파트에서 배운 여러가지 불균형 데이터 처리 기법을 적용해보고 그 결과를 살펴보겠습니다.

2. Final Project

a. 실전 머신러닝 I

- i. 출제자 : 이경록 (브레인크루 CEO / Data Scientist)
- ii. 기한 : 2020. 11. 09 - 2020. 11. 23
- iii. 프로젝트 명 : 대출 상환 예측 (Home Credit Default Risk)
- iv. 프로젝트 설명 :
많은 사람들은 신용 기록이 불충분하거나 존재하지 않아 대출을 받기 위해 애를 씁니다. 주택 신용 기관(Home Credit)은 안전한 차용을 제공하고자 고객의 상환 능력을 미리 예측할 필요가 있고, 이를 위해 전화 및 거래 정보를 포함한 다양한 데이터를 사용하고 있습니다. 여러분들은 지금까지 배운 다양한 통계 분석 기법과 머신러닝 모델을 활용하여 고객의 부채 상환 능력에 대해 예측을 진행하게 됩니다.

b. 실전 머신러닝 II

- i. 출제자 : 이경록 (브레인크루 CEO / Data Scientist)
- ii. 기한 : 2020. 11. 23 - 2020. 12. 06
- iii. 프로젝트 명 : 집값 예측 - 심화 (House Prices : Advanced Regression Techniques)
- iv. 프로젝트 설명 :
많은 사람들이 원하는 집의 조건은 무엇일까요? 집을 묘사할 수 있는 데이터는 무수히 많습니다. 집의 천장 높이가 될 수도 있고 각종 구조물일 수도 있겠습니다. 본 프로젝트에서는 79개의 다양한 '집' 관련 데이터를 기반으로, 사람들이 원하는 집을 상상해보고 이를 통해 최종 집값을 예측하게 됩니다. 데이터의 이상치, 결측치를 처리해보는 것은 물론, 회귀분석부터 다양한 모델들의 고급 머신러닝 기법, 앙상블 학습까지 전 범위의 내용을 통합하여 모델의 성능을 끌어올려봅니다.

1. 8회 이상 출석 체크
 - a. 총 12주 동안 12번의 출석 체크를 진행하고 2/3 이상(8회 이상) 출석 인증을 해야합니다.
 - b. 출석 체크는 퀴즈 통과 여부를 통해 인증합니다.
 - c. 매주 수업이 끝날 때, 10점 만점의 문제가 출제 됩니다.
 - d. 5점 이상 득할 시, 1회 출석이 인정됩니다.
2. 4회 이상 Mini Project 제출
 - a. '머신러닝을 위한 수학, 머신러닝의 시작-회귀분석, 머신러닝 기본모형, 앙상블, 군집분석&불균형데이터 다루기' 5개의 파트에서 총 6개의 Mini Project가 출제 됩니다.
 - b. 6개의 프로젝트 중 4개 이상의 Mini Project를 제출 하여야 합니다.
 - c. 제출된 Mini Project는 리뷰어를 통해 피드백을 진행합니다.
3. 2회 Final Project 통과
 - a. 실전 프로젝트 파트에서 총 2개의 Final Project가 출제 됩니다.
 - b. 모든 수업을 종료된 이후 4주의 Final Project 출제 기간이 주어집니다.
 - c. 리뷰어를 통해 피드백이 진행되고, 리뷰어는 공개된 기준표에 맞추어 Project 통과 여부를 결정하며 피드백을 진행합니다.
※ 리뷰어의 피드백은 프로젝트 제출일부터, 최대 3일 이내에 진행됩니다.
 - d. 리뷰어의 '통과'가 이루어진 수강생 중 출석체크, Mini Project 제출 수료 기준을 만족한 수강생에게 Byte Degree가 발급됩니다.

10. Course Curriculum

파트	상세 클립 명	클립명 시간
머신러닝 개요	Machine Learning의 개념	11:22
	지도학습과 비지도학습	7:10
	Machine Learning의 종류	15:40
	모형의 적합성 평가 및 실험설계 - 1	13:43
	모형의 적합성 평가 및 실험설계 - 2	14:07
머신러닝을 위한 수학	과적합(Overfitting)이란	11:55
	수학적 개념 이해 - 통계학 개념, 자료의 종류, 자료의 요약	10:35
	수학적 개념 이해 - 기술통계량, 추정량	16:39
	수학적 개념 이해 - 확률의 기초	15:29
	수학적 개념 이해 - 확률 변수, 확률 분포	11:37
	수학적 개념 이해 - 이산형 확률 분포, 연속형 확률 분포	13:49
	수학적 개념 이해 - 추정, 추론	17:18
	수학적 개념 이해 - 유의확률	7:30
	수학적 개념 이해 - 검정통계량의 분포	8:36
	수학적 개념 이해 - 미분의 개념	15:48
	수학적 개념 이해 - Likelihood, 확률분포함수	9:32
	수학적 개념 이해 - MLE 개념	7:56
	수학적 개념 이해 - MLE 계산방법	12:52
	수학적 개념 이해 - Matrix 정의 및 성질	16:26
	수학적 개념 이해 - Matrix 미분 기초	12:30
머신러닝의 시작, 회귀분석	수학적 개념 이해 - 주요 Matrix 미분 결과 및 증명	10:44
	수학적 개념 이해 - 주요 Matrix 미분 결과 정리	10:44
	수학적 개념 이해 - Matrix 미분의 회귀분석에서의 활용	17:03
	회귀분석이란	7:19
	회귀계수를 추정하는 방법	13:13
	회귀계수의 의미	12:29
	회귀계수에 대한 검정	9:37

단순선형회귀 적합 및 해석 (실습)	21:59
다중선형회귀분석에 대한 개념	10:02
다중선형회귀분석의 회귀계수	10:18
다중회귀분석 적합 및 단순선형회귀와의 비교 (실습)	13:34
다중공선성이란 무엇인가	8:24
다중공선성 진단 방법	10:23
회귀모델의 성능지표	7:09
모델의 성능지표	13:09
다중회귀 모델 해석 및 다중공선성진단 (실습) - 1	16:33
다중회귀 모델 해석 및 다중공선성진단 (실습) - 2	16:41
변수선택법	8:50
교호작용	6:32
다중회귀모형에 대한 검증	13:07
다항회귀분석	6:21
다중선형회귀분석 - 변수선택법을 통한 다중회귀분석 (실습) - 1	23:28
다중선형회귀분석 - 변수선택법을 통한 다중회귀분석 (실습) - 2	23:25
로지스틱 회귀분석 이란	6:54
로지스틱 회귀계수 추정법과 해석	8:17
로지스틱 회귀분석 - 로지스틱 회귀모델 적합 및 해석 (실습)	23:34
회귀계수를 축소하는 이유	6:56
회귀계수 축소법의 종류와 차이점 - 1	8:39
회귀계수 축소법의 종류와 차이점 - 2	13:27
회귀계수 축소법 - Lasso, Ridge 적합 및 로지스틱회귀와 비교 (실습)	16:27
Feature selection 정리	11:11
머신러닝 기본모형	
나이브 베이즈 - 01. Naïve bayes classifier 배경	8:01
나이브 베이즈 - 02. 수학적 개념 이해 - 조건부 확률, Bayes 정리	13:35
나이브 베이즈 - 03. Naïve Bayes classifier 정의	11:46
나이브 베이즈 - 04. Naïve bayes 종류 및 이해	7:51
나이브 베이즈 - 05. Naïve bayes 프로그램 실습	14:04
KNN - 01. k-Nearest Neighbor 배경	10:50
KNN - 02. k-Nearest Neighbor 모델 정의 및 추정	5:38
KNN - 03. 수학적 개념 이해 - Cross-validation	9:38
KNN - 04. KNN의 심화적 이해	14:05
KNN - 05. KNN 프로그램 실습	7:47
SVM - 01. Support Vector Machine 배경	6:05
SVM - 02. 수학적 개념 이해 - Decision boundary, decision rule	3:35
SVM - 03. 수학적 개념 이해 - Lagrange multiplier	8:00
SVM - 04. SVM 정의	12:20
SVM - 05. SVM 계산 및 Support vectors 이해	8:53
SVM - 06. SVM 심화적 이해	7:18
SVM - 07. One-Class SVM	7:42
SVM - 08. Support Vector Regression	6:34
SVM - 09. SVM 프로그램 실습	8:33
의사결정나무 - 01. 의사결정나무 배경	10:21
의사결정나무 - 02. 수학적 개념 이해 - 엔트로피	10:52
의사결정나무 - 03. 수학적 개념 이해 - Information Gain	11:22

	의결정나무 - 04. Classification Tree 정의 및 계산	8:43	
	의사결정나무 - 05. Regression Tree 정의 및 계산	6:49	
	의사결정나무 - 07. 의사결정나무 프로그램 실습 - 1	24:51:00	
	의사결정나무 - 08. 의사결정나무 프로그램 실습 - 2	19:08	
앙상블	Ensemble의 개념 - 01. Ensemble이란 - 1	6:32	
	Ensemble의 개념 - 02. Ensemble이란 - 2	8:07	
	Ensemble의 개념 - 03. Ensemble의 종류	7:02	
	Bagging - 01. Bagging	7:36	
	Bagging - 02. Bagging 실습 (for문을 이용하여 bagging알고리즘 작성)	24:33:00	
	Bagging - 03. Bagging 실습 (패키지를 이용한 bagging 실습)	17:07	
	RandomForest - 01. RandomForest	7:06	
	RandomForest - 02. RandomForest 실습	14:19	
	Boosting - 01. Boosting	5:58	
	Boosting - 02. Gradient Boosting	9:21	
	Boosting - 03. GradientBoosting의 종류	14:30	
	Boosting - 04. Boosting 실습 (Adaboost 실습)	8:19	
	Boosting - 05. Boosting 실습 (Gradient Boosting 실습)	19:39	
	Stacking - 01. Stacking	10:27	
	Stacking - 02. Ensemble의 Ensemble	7:20	
	Stacking - 03. Ensemble의 Ensemble 실습	10:41	
	Stacking - 04. Ensemble기법 review	8:48	
	Shap value - 01. 중요 변수 추출 방법	11:27	
	Shap value - 02. Shap value 소개	14:15	
	Shap value - 03. Shap value 실습 (Shap value를 통한 모형 해석)	11:31	
	Shap value - 04. DS분야에서 Tree기반 모델이 쓰이는 이유	12:09	
	군집분석 & 불균형데이터 다루기	K-means - 01. Clustering이란	6:23
		K-means - 02. K-means clustering	5:57
		K-means - 03. 최적의 K를 찾는법	10:48
K-means - 04. K-medoid clustering		5:39	
K-means - 05. K-medoid clustering 실습		13:59	
Hierarchical clustering - 01. Hierarchical clustering		7:38	
Hierarchical clustering - 02. Hierarchical clustering 실습		7:41	
DBSCAN clustering - 01. DBSCAN clustering		12:07	
DBSCAN clustering - 02. DBSCAN 실습		12:49	
Class Imbalanced problem이란		14:13	
Class Imbalanced problem을 해결하기 위한 기법 소개		10:45	
Oversampling기법		8:55	
SMOTE변형 기법들		10:36	
SMOTE 실습 A - SMOTE 적용		23:23	
SMOTE 실습 B - 다양한 모델 적용		12:13	
SMOTE 실습 C - Ensemble 모델 적용		25:47:00	
Undersampling기법	6:28		
Hybrid resampling기법	15:26		

(끝)