

K-Digital Credit

# 왕초보로 시작하는 머신러닝 프로젝트 완성

- » 데이터 사이언스를 위한
- » 최고의 시작

## I | 과정 소개

- ▶ 과정이름 : 왕초보로 시작하는 머신러닝 프로젝트 완성
- ▶ 선수과목 : 없음
- ▶ 수강비용 : 내일배움카드 활용시 25,000원 (수료시 100%환급)
- ▶ 지원대상 : 4년제 3학년 / 전문대 1학년부터 34세 이하 청년,  
혹은 만 35세 이상 55세 미만 구직자

## II | 학습 내용

### 1. 제공 학습 콘텐츠

- 영상강의 (정규강의 14시간)
- 실습용 온라인 워크북 (<https://school.coding-x.com/>)
- 실습 워크 Sheet, 해설 코드, 질문 답변 및 과제 피드백

### 2. 코스 소개

코스 이름	코스 소개
AI란 무엇인가?	- AI에 대한 기본 개념과 입문을 위한 학습
인공지능을 위한 파이썬	- AI를 다루기 위한 기초 언어 파이썬 학습
데이터 분석 입문	- 데이터 분석에 대한 기초 학습 진행
데이터 분석 활용	- 클라우드 분석환경에서 데이터 분석 활용
머신러닝 알고리즘	- 머신러닝의 다양한 알고리즘에 대한 학습
머신러닝의 흐름	- 준비단계부터 머신러닝 모델의 완성까지 학습
프로젝트 1	- 주가 예측을 바탕으로 수익률을 계산해 보는 프로젝트
프로젝트 2	- 대학교에 대한 데이터를 사용하여 군집분석하는 프로젝트
프로젝트 3	- 통신사 고객들의 이탈 여부를 예측하는 프로젝트

### III 세부 학습 내용

챕터 이름	학습 목표
The! 인공지능 인공지능 알아보기	<ol style="list-style-type: none"> <li>1. 인공지능을 이해하는 방법에 대한 학습한다.</li> <li>2. 인공지능의 성격과 현황에 대해 이해할 수 있다.</li> <li>3. 우리 생활속에 활용되고 있는 인공지능 사례를 이해할 수 있다.</li> </ol>
The! 인공지능 인공지능 학습 시작하기	<ol style="list-style-type: none"> <li>1. AI의 학습 요소와 필요성에 대해 이해할 수 있다.</li> <li>2. AI를 공부하기 위해 파이썬을 사용해야하는 이유를 이해할 수 있다.</li> <li>3. 라이브러리의 개념과 파이썬의 대표적인 라이브러리를 알아본다.</li> </ol>
인공지능을 위한 파이썬 파이썬의 시작	<ol style="list-style-type: none"> <li>1. 컴퓨터 언어에 대해 이해할 수 있다.</li> <li>2. 파이썬의 탄생 배경과 특징에 대해 이해할 수 있다.</li> <li>3. 파이썬을 학습하는 방법에 대해 이해하고, 스스로 목표를 세울 수 있다.</li> </ol>
인공지능을 위한 파이썬 파이썬 실습환경	<ol style="list-style-type: none"> <li>1. 코딩엑스의 파이썬 실습환경에 대해 이해할 수 있다.</li> <li>2. 파이썬의 주석에 대해 이해하고 사용할 수 있다.</li> </ol>
인공지능을 위한 파이썬 파이썬의 기초	<ol style="list-style-type: none"> <li>1. print 함수와 옵션에 대해 이해하고 사용할 수 있다.</li> <li>2. 할당의 개념을 이해하고 사용할 수 있다.</li> <li>3. 문자열을 표현하는 여러가지 방법에 대해 이해할 수 있다.</li> </ol>
인공지능을 위한 파이썬 수학? 파이썬!	<ol style="list-style-type: none"> <li>1. 파이썬을 사용하여 사칙연산을 할 수 있다.</li> <li>2. 숫자를 다루는 함수에 대해 이해하고 사용할 수 있다.</li> <li>3. 문자열의 덧셈과 곱셈을 이해하고 활용할 수 있다.</li> </ol>
인공지능을 위한 파이썬 변수와 데이터	<ol style="list-style-type: none"> <li>1. 변수의 이름을 짓는 규칙에 대해 이해할 수 있다.</li> <li>2. 데이터 타입의 종류와 특징에 대해 이해할 수 있다.</li> </ol>
인공지능을 위한 파이썬 문자열 포매팅	<ol style="list-style-type: none"> <li>1. format 함수의 사용 이유를 이해하고 활용할 수 있다.</li> <li>2. 식지정자를 통해 포매팅 하는 방법을 이해하고 활용할 수 있다.</li> </ol>
인공지능을 위한 파이썬 컬렉션 데이터	<ol style="list-style-type: none"> <li>1. 리스트, 튜플, 집합, 딕셔너리 자료구조에 대해 이해하고 사용할 수 있다.</li> <li>2. 리스트 자료구조를 수정할 수 있다.</li> </ol>
인공지능을 위한 파이썬 조건과 반복	<ol style="list-style-type: none"> <li>1. 조건문의 원리를 이해하고 논리 구조를 작성할 수 있다.</li> <li>2. 반복문에 대해 이해하고 활용할 수 있다.</li> </ol>
인공지능을 위한 파이썬 함수에 대해	<ol style="list-style-type: none"> <li>1. 여러가지 함수의 형태에 대해 이해하고 사용할 수 있다.</li> <li>2. 사용자 정의 함수의 다양한 기능들을 익히고 활용할 수 있다.</li> </ol>
데이터 분석 입문 데이터를 활용하는 방법	<ol style="list-style-type: none"> <li>1. 데이터 활용 이유에 대해 설명 할 수 있다.</li> <li>2. 데이터의 다양한 종류를 구분할 수 있다.</li> </ol>
데이터 분석 입문 데이터 구조 이해하기	<ol style="list-style-type: none"> <li>1. 정형데이터의 기본 구조를 이해하고 행과 열의 개념을 설명할 수 있다.</li> <li>2. 데이터의 다양한 속성을 확인하고 구분 할 수 있다.</li> <li>3. 실제로 사용되는 데이터의 형태를 이해하고 데이터의 특징을 설명할 수 있다.</li> </ol>
데이터 분석 입문 데이터를 다루는 도구	<ol style="list-style-type: none"> <li>1. 데이터 분석에 파이썬 프로그램을 이용하는 이유에 대해 이해할 수 있다.</li> <li>2. 파이썬의 라이브러리 개념에 대해 이해할 수 있다.</li> <li>3. 데이터 분석에 자주 사용되는 두 가지 라이브러리에 대해 알아본다.</li> </ol>

챕터 이름	학습 목표
<b>데이터 분석 입문</b> <b>Pandas에 대해 알아보자</b>	1. pandas 라이브러리를 import하고 데이터를 저장하고 불러올 수 있다. 2. pandas를 통해 데이터의 구조를 파악하고 선택하여 확인할 수 있다. 3. pandas를 통해 데이터를 복사, 생성, 삭제하고 특정 조건으로 탐색할 수 있다.
<b>데이터 분석 입문</b> <b>이상한 데이터?</b>	1. 결측치의 의미를 이해하고 결측치의 위치를 찾을 수 있다. 2. 이상치의 의미를 이해하고 이상치의 위치를 확인할 수 있다. 3. 결측치와 이상치를 제거할 수 있다.
<b>데이터 분석 입문</b> <b>데이터 살펴보기</b>	1. 범주형 데이터를 탐색할 수 있다. 2. 수치형 데이터를 탐색할 수 있다. 3. 간단한 시각화를 통해 데이터를 탐색할 수 있다.
<b>데이터 분석 입문</b> <b>상관관계 파악하기</b>	1. 상관관계에 대해 이해할 수 있다. 2. 상관분석을 하고 결과를 해석할 수 있다.
<b>데이터 분석 입문</b> <b>그룹별 연산</b>	1. 그룹별 연산에 대해 이해할 수 있다. 2. groupby 함수를 사용하여 그룹별 연산을 할 수 있다. 3. groupby 함수를 사용하여 여러 개의 연산을 동시에 처리할 수 있다.
<b>데이터 분석 활용</b> <b>편리한 작업 환경</b>	1. 구글 코랩 Colaboratory에 대해 이해할 수 있다. 2. 코랩에 접속하여 사용 환경을 준비할 수 있다. 3. 코랩의 기본 작동 방법을 이해할 수 있다.
<b>데이터 분석 활용</b> <b>탐색적 데이터 분석</b>	1. 데이터 분석에 필요한 패키지를 불러올 수 있다. 2. iris 데이터를 불러오고 탐색할 수 있다. 3. 분석에 필요한 데이터를 선별하고 결측치를 확인할 수 있다.
<b>데이터 분석 활용</b> <b>탐색적 데이터 분석2</b>	1. 수치형 데이터의 기초통계를 확인하고 히스토그램을 그릴 수 있다. 2. 범주형 데이터의 빈도와 비율을 확인하고 파이 그래프로 나타낼 수 있다. 3. 특성별 산점도를 그리고 상관분석을 통해 결과를 해석할 수 있다.
<b>데이터 분석 활용</b> <b>그룹별 연산 적용</b>	1. 범주형 데이터와 수치형 데이터를 함께 활용하여 의미를 도출할 수 있다. 2. 그룹별로 다양한 연산을 묶어서 처리 할 수 있다.
<b>데이터 분석 활용</b> <b>데이터 변환</b>	1. 범주형 데이터를 수치형 데이터로 변환할 수 있다. 2. 수치형 데이터를 범주형 데이터로 변환할 수 있다. 3. 새로운 데이터 특성을 생성할 수 있다.
<b>데이터 분석 활용</b> <b>데이터 시각화1</b>	1. 시각화의 중요성과 그래프의 종류별 사용 방법을 이해할 수 있다. 2. 박스 플롯을 그리고 해석할 수 있다. 3. 선그래프와 히트맵 그래프를 작성하고 해석할 수 있다.
<b>데이터 분석 활용</b> <b>데이터 시각화2</b>	1. 그래프 제목과 x축, y축 이름을 추가할 수 있다. 2. 그래프에 색상을 지정할 수 있다. 3. 막대 그래프의 디자인 옵션을 다룰 수 있다.
<b>데이터 분석 활용</b> <b>분석 과정 리뷰</b>	1. 새로운 데이터를 통해 데이터 준비 및 정리 과정을 실습할 수 있다. 2. 새로운 데이터를 통해 데이터 탐색 과정을 실습할 수 있다. 3. 새로운 데이터를 통해 데이터 시각화 과정을 실습할 수 있다.
<b>머신러닝 알고리즘</b> <b>머신러닝 개요</b>	1. 머신러닝의 개념을 이해할 수 있다. 2. 머신러닝의 종류를 이해할 수 있다. 3. 머신러닝의 프로세스에 대해 이해할 수 있다.
<b>머신러닝 알고리즘</b> <b>K-NN(최근접 이웃법)</b>	1. K-최근접 이웃법의 개념에 대해 이해할 수 있다. 2. 유클리디안 거리를 이해하고 계산할 수 있다. 3. K-NN 모델을 구현할 수 있다.
<b>머신러닝 알고리즘</b> <b>의사결정 나무</b>	1. 의사결정나무의 개념을 이해할 수 있다. 2. 의사결정나무가 조건을 만드는 기준을 이해할 수 있다. 3. 의사결정나무 모델을 구현할 수 있다.

챕터 이름	학습 목표
<b>머신러닝 알고리즘</b> <b>앙상블과 랜덤포레스트</b>	1. 앙상블 개념에 대해 이해할 수 있다. 2. 랜덤포레스트 개념에 대해 이해할 수 있다. 3. 랜덤포레스트 모델을 구현할 수 있다.
<b>머신러닝 알고리즘</b> <b>K-means 군집분석</b>	1. 군집분석의 개념을 이해할 수 있다. 2. K-means 군집분석 방법에 대해 이해할 수 있다. 3. K-means 모델을 구현할 수 있다.
<b>머신러닝 알고리즘</b> <b>연관규칙분석</b>	1. 연관규칙분석의 개념을 이해할 수 있다. 2. 연관규칙분석의 측정지표에 대해 이해할 수 있다. 3. 연관규칙분석 모델을 구현할 수 있다.
<b>머신러닝 알고리즘</b> <b>신경망</b>	1. 신경망 모형의 개념을 이해할 수 있다. 2. 퍼셉트론의 작동 원리와 딥러닝의 개념을 이해할 수 있다. 3. 신경망 모델을 구현할 수 있다.
<b>머신러닝의 흐름</b> <b>데이터 준비와 분석 목표</b>	1. 분석 데이터의 특성을 이해할 수 있다. 2. 와인 데이터를 활용하여 분석의 목표를 설정할 수 있다.
<b>머신러닝의 흐름</b> <b>데이터 나누기</b>	1. 데이터를 분할하는 이유를 이해할 수 있다. 2. 필요에 따라 데이터를 분할할 수 있다.
<b>머신러닝의 흐름</b> <b>데이터 스케일링</b>	1. 데이터 스케일링의 개념을 이해할 수 있다. 2. 데이터를 스케일링 하여 사용할 수 있다.
<b>머신러닝의 흐름</b> <b>클래스 불균형</b>	1. 클래스 불균형에 대해 이해할 수 있다. 2. 오버샘플링을 통해 클래스 불균형을 처리할 수 있다.
<b>머신러닝의 흐름</b> <b>하이퍼파라미터 탐색</b>	1. 하이퍼파라미터와 그리드 서치의 개념에 대해 이해할 수 있다. 2. 그리드서치를 이용하여 적절한 하이퍼 파라미터를 찾을 수 있다.
<b>머신러닝의 흐름</b> <b>모델 성능 평가</b>	1. 오차행렬과 모델 성능 평가 지표에 대해 이해할 수 있다. 2. 성능평가 지표와 오차행렬을 계산할 수 있다.
<b>프로젝트 1차(주가예측)</b> <b>이론 강의 및 방향</b>	1. 기본적인 이진분류 문제에 분류 모델을 적용하고 파라미터를 튜닝할 수 있다. 2. 분류 결과를 문제의 목적에 맞춰 적절히 해석할 수 있다.
<b>프로젝트 1차(주가예측)</b> <b>해설 강의</b>	1. 기본적인 이진분류 문제에 분류 모델을 적용하고 파라미터를 튜닝할 수 있다. 2. 분류 결과를 문제의 목적에 맞춰 적절히 해석할 수 있다.
<b>프로젝트 2차(대학 군집분석)</b> <b>이론 강의 및 방향</b>	1. 군집분석에 필요한 전처리와 실제 분석을 수행할 수 있다. 2. 군집분석의 결과를 적절히 해석할 수 있다.
<b>프로젝트 2차(대학 군집분석)</b> <b>해설 강의</b>	1. 군집분석에 필요한 전처리와 실제 분석을 수행할 수 있다. 2. 군집분석의 결과를 적절히 해석할 수 있다.
<b>프로젝트 3차</b> <b>(통신사 가입자 이탈)</b> <b>이론 강의 및 방향</b>	1. 복잡한 데이터에 대한 전처리와 탐색적 데이터 분석을 수행할 수 있다. 2. 성능지표를 고안하고 성능지표에 맞는 모델 선택을 수행할 수 있다.
<b>프로젝트 3차</b> <b>(통신사 가입자 이탈)</b> <b>해설 강의</b>	1. 복잡한 데이터에 대한 전처리와 탐색적 데이터 분석을 수행할 수 있다. 2. 성능지표를 고안하고 성능지표에 맞는 모델 선택을 수행할 수 있다.

## IV 주차별 학습 계획

- ▷ 실습과 학습 과제를 포함하여 주별 2시간 학습시 12주 소요
- ▷ 계획과 무관하게 선행 학습 가능

<p><b>1주차</b></p> <p><b>1 - 6 차시</b></p> <ul style="list-style-type: none"> <li>- 오리엔테이션</li> <li>- 인공지능 기초개념</li> <li>- 인공지능 학습방법</li> <li>- 파이썬 소개</li> <li>- 파이썬 실습환경</li> <li>- 파이썬의 기초</li> <li>- 파이썬의 연산</li> </ul>	<p><b>2주차</b></p> <p><b>7 - 11 차시</b></p> <ul style="list-style-type: none"> <li>- 변수와 데이터</li> <li>- 문자열 포매팅</li> <li>- 컬렉션 데이터</li> <li>- 조건문과 반복문</li> <li>- 함수의 사용</li> </ul>	<p><b>3주차</b></p> <p><b>12 - 15 차시</b></p> <ul style="list-style-type: none"> <li>- 데이터 활용 방법</li> <li>- 데이터 구조 이해</li> <li>- 데이터를 다루는 도구</li> <li>- Pandas 사용법</li> </ul>	<p><b>4주차</b></p> <p><b>16 - 19 차시</b></p> <ul style="list-style-type: none"> <li>- 결측치와 이상치</li> <li>- 데이터 살펴보기</li> <li>- 상관관계 파악</li> <li>- 그룹별 연산</li> </ul>
<p><b>5주차</b></p> <p><b>20 - 24 차시</b></p> <ul style="list-style-type: none"> <li>- 클라우드 분석환경</li> <li>- 탐색적 데이터 분석</li> <li>- 그룹별 연산 적용</li> <li>- 데이터 변환</li> </ul>	<p><b>6주차</b></p> <p><b>25 - 27 차시</b></p> <ul style="list-style-type: none"> <li>- 데이터 시각화</li> <li>- 시각화 디자인 옵션</li> <li>- 분석과정 리뷰</li> </ul>	<p><b>7주차</b></p> <p><b>28 - 30 차시</b></p> <ul style="list-style-type: none"> <li>- 머신러닝 개요</li> <li>- K-NN 최근접 이웃법</li> <li>- 의사결정나무</li> </ul>	<p><b>8주차</b></p> <p><b>31 - 34 차시</b></p> <ul style="list-style-type: none"> <li>- 앙상블 랜덤포레스트</li> <li>- K-means 군집분석</li> <li>- 연관규칙분석</li> <li>- 신경망 분석</li> </ul>
<p><b>9주차</b></p> <p><b>35 - 41차시</b></p> <ul style="list-style-type: none"> <li>- 분석목표 설정</li> <li>- 데이터 분할</li> <li>- 클래스 불균형</li> <li>- 하이퍼파라미터 탐색</li> <li>- 모델 성능 평가</li> </ul>	<p><b>10주차</b></p> <p><b>42 - 43차시</b></p> <ul style="list-style-type: none"> <li>- 1차 프로젝트 강의</li> <li>- 1차 프로젝트 수행</li> </ul>	<p><b>11주차</b></p> <p><b>44 - 45차시</b></p> <ul style="list-style-type: none"> <li>- 2차 프로젝트 강의</li> <li>- 2차 프로젝트 수행</li> </ul>	<p><b>12주차</b></p> <p><b>46 차시</b></p> <ul style="list-style-type: none"> <li>- 3차 프로젝트 강의</li> <li>- 3차 프로젝트 수행</li> </ul>