인공지능 심화 학습 보고서

|  |  |  |  |
| --- | --- | --- | --- |
| **작성일자** | 2023.01.30 | **작성자** | 박현준 |
| **과제명** | 코로나 데이터를 이용한 사망예측 | | |
| **학습동기** | # 코로나 dataset으로 선정한 이유   * 최근에 코로나에 걸렸을 때 실제로 사망하신 중증코로나 환자 분들은 기저질환과 어떤 연관성이 있을까 궁금하였고 이를 딥러닝으로 분석하면 좋겠다라는 생각이 들어서 코로나 데이터를 찾아서 사용하였다. | | |
| **학습 알고리즘의 개념 정리** | | | |
| # 선형회귀   * 임의의 직선 y = ax + b와 각각의 데이터 요소와의 직선 거리를 이용하여 예측하는 알고리즘 * 그러나 실제로 알고리즘은 거리를 제곱하여 사용한다.   그 이유는 무엇일까?  절댓값을 사용하면 불연속 함수가 되기 때문에 제곱을 한다  그러나 데이터의 양이 많아지고 값이 커지면 보기 힘들기 때문에  루트를 사용하여 제곱근을 이용한다고 한다.  A line that models a set of data | | | |
| **연구내용** | | | |
| # 개발내용 정리   * COVID-19.csv 파일을 이용함 * 각 열의 항목     을 이용하여 사망여부 판단   * 사용한 라이브러리 * numpy : Numpy는 C언어로 구현된 파이썬 라이브러리로써, 고성능의 수치계산을 위해 제작됨. * Tensorflow : 구글(Google)에서 만든, 딥러닝 프로그램을 쉽게 구현할 수 있도록 다양한 기능을 제공해주는 라이브러리 * Keras : [파이썬](https://ko.wikipedia.org/wiki/%ED%8C%8C%EC%9D%B4%EC%8D%AC)으로 작성된 [오픈 소스](https://ko.wikipedia.org/wiki/%EC%98%A4%ED%94%88_%EC%86%8C%EC%8A%A4_%EC%86%8C%ED%94%84%ED%8A%B8%EC%9B%A8%EC%96%B4) [신경망](https://ko.wikipedia.org/wiki/%EC%9D%B8%EA%B3%B5%EC%8B%A0%EA%B2%BD%EB%A7%9D) 라이브러리 * Sklearn : 사이킷런은 파이썬에서 머신러닝 분석을 할 때 유용하게 사용할 수 있는 라이브러리   # 데이터 분석에 필요한 라이브러리 import  텍스트이(가) 표시된 사진  자동 생성된 설명  텍스트이(가) 표시된 사진  자동 생성된 설명  # Google Colab에 업로드 된 CSV 데이터 파일 불러오기   * 사망여부를 참과 거짓으로 나타내기 위해서 class 라는 열을 마지막 열에 추가하였다. 이때 class 라는 열은 초기에 1이라는 데이터로 채워져있는 상태인데 1은 사망을 하지 않았음을 의미한다 * DATE\_DIED(사망날짜) 열을 찾아서 값이 ‘9999-99-99’라는 값이 들어있는 행만 제외하고 나머지 행의 class 부분에 0을 대입한다. * 이 작업을 함으로써 사망여부를 참, 거짓 두가지로 나타낼 수 있게 되었고 이진분류를 활용한 학습을 할 수 있게 되었다   텍스트이(가) 표시된 사진  자동 생성된 설명  # 전체 데이터의 0.3(30%)을 검증하기 위한 데이터로 정함   * 과적합을 피하고 편향을 제거한 데이터로 모델 성능을 평가하기 위해서다     # 활성화 함수를 사용하여 데이터 학습     * 학습원리      * Sigmoid * 이진분류를 위한 활성화 함수 * 구분 하려는 데이터가 참 또는 거짓(2종류)일 때 사용   # 반복학습     * Epochs : 반복횟수(즉, 약 73만개의 검증데이터를 학습절차를 100번 반복)   # 딥러닝 결과      93% 의 정확도라는 결과가 나옴 | | | |
| **소감 및 향후 계획** | | | |
| # 반성 및 소감   * 처음에 인공지능에 흥미를 느끼지 못하였고 인공지능을 구현하는 과정을 이해하는 것이 어려웠다 * 학습모델을 설정할 때 나오는 여러가지 함수 (예 : sigmoid)를 이해하기 위해서 미분의 개념을 공부하는 것이 어려웠다. 왜냐하면 아직 미분에 대해서 배우지 않았기 때문이다. 그래서 고등학교 학년이 올라가고 미분을 배우면 조금 더 인공지능을 쉽게 접할 수 있을 것 같았다 * 학습 model을 설정하고 구현할 때 numpy를 이용해서 설정을 했어야 하지만 그 점을 잘 몰랐기 때문에 한참을 해매었다. 그러나 결국 해결해서 뿌듯하였다.   # 향후 계획   * 개인적으로 인공지능에 관한 공부를 더 하여서 여러가지 인공지능 모델을 만들어 보고 싶다 * 학교에서 딥러닝을 공부해서 경진대회를 나갈려다가 공부가 어렵기도 하였고 여러가지 일정이 겹쳐서 못 나갔었는데 딥러닝을 경험해 봄으로써 여러가지 대외활동에 참여하여 사용할 것이다. * 파이썬 공부를 열심히 할 것이다 | | | |