

인텔리전트 데이터베이스 서비스를 위한 지능형 API 설계 및 구현

임종태, 함동호, 김남영, 김윤아, 이소민, 신보경, 이현병, 유재수*

충북대학교

jtlim@cbnu.ac.kr, ii8858@naver.com, minstrel68@naver.com, rud5356@naver.com,
{somin, sbk02, lhb, yjs}@cbnu.ac.kr

Design and Implementation of Intelligent APIs for Intelligent Database Services

Jongtae Lim, Dongho Ham, Nameyoung Kim, Yuna Kim1, Somin Lee, Bokyoung Shin,

Hyeonbyeong Lee, Jaesoo Yoo

Chungbuk National University

요약

최근 빅데이터를 분석하는 기술이 중요하게 연구됨에 따라 기계 학습이 이슈가 되고 있다. 기계 학습을 이용하여 서비스를 개발하기 위해서는 서비스에 활용할 학습 모델을 생성해야 하는데 이는 기계 학습에 대한 전문성이 요구되기 때문에 기계 학습을 활용한 지능형 서비스 개발은 쉽지 않은 것이 현실이다. 본 논문에서는 인텔리전트 데이터베이스를 활용한 서비스를 위한 지능형 API를 설계한다. 제안하는 API는 모델 생성 지원 API와 서비스 개발 지원 API로 구성된다. 모델 생성 지원 API는 기계 학습 모델 생성을 위해 필요한 기능을 실행하기 위한 API이다. 그리고 서비스 개발 지원 API는 인텔리전트 데이터베이스에 저장된 데이터, 학습된 모델 등을 활용하여 서비스를 개발하기 위해 필요한 기능을 실행하기 위한 API이다.

I. 서론

최근 빅데이터 분석 방법 중에 하나로 기계 학습이 중요하게 연구되고 있다. 기계 학습은 스팸메일필터와 같은 간단한 서비스부터 인공지능 비서와 같은 복잡한 서비스까지 활용할 수 있는 중요한 기술이다. 하지만 기계 학습을 이용하여 서비스를 개발하기 위해서는 기계 학습을 이용하여 학습 모델을 생성하기 위한 전문적인 지식이 필요하다. 기계 학습을 쉽게 수행하기 위한 다양한 도구들이 소개되었지만 여전히 기계 학습은 높은 수준의 전문성을 필요로 한다[1]. 따라서 기계 학습을 활용한 서비스를 제공하기 위해 자율 기계학습을 수행하고, 결과를 분석하여 지식을 축적할 수 있는 인텔리전트 데이터베이스가 요구된다. 최근 구글(Google), 파나소니(Panasonic) 등의 기업들을 중심으로 자율 기계학습 플랫폼들이 발표됐다[2]. 본 논문에서는 인텔리전트 데이터베이스에 기반을 둔 자율 기계 학습을 수행하기 위하여 지능형 API를 설계하고 구현한다. 제안하는 API는 모델 생성 지원 API와 서비스 개발 지원 API로 구성된다. 기계 학습 개발자와 서비스 개발자가 본 논문에서 제안하는 인텔리전트 데이터베이스 시스템에 기반을 둔 API를 활용하여 모델을 생성하거나 등록, 검색하여 서비스에 활용할 모델을 선정한다. 또한 저장된 모델을 이용하여 추론을 수행하여 서비스에 활용하는 전체 과정을 수행할 수 있다.

II. 지능형 API 설계 및 구현

1. 지능형 API 설계

그림 1은 인텔리전트 데이터베이스 사용자를 위한 지능형 API 설계 목록을 보여준다. 제안하는 API는 모델 생성 지원 API와 서비스 개발 지원 API로 구성된다. 인텔리전트 데이터베이스 사용자는 크게 기계학습 개발자, 어플리케이션 개발자, 서비스 사용자로 구분된다. 기계 학습 개발자는 기계 학습에 대한 전문가로 지능형 서비스에서 활용하기 위한 학습 모델

을 생성하는 역할을 수행한다. 기계 학습 개발자는 서비스 목적에 부합하는 기계 학습 알고리즘을 설정한다. 기계 학습 개발자는 학습대상 데이터의 가공 및 전처리를 수행하여 기계 학습을 수행하고 도출된 학습 모델의 성능 측정 및 재학습을 통한 성능 향상 등 기계 학습과 관련된 다양한 역할을 수행한다. 어플리케이션 개발자는 특정 목적을 가진 지능형 서비스를 개발하는 개발자로 기계 학습 개발자의 역할을 함께 수행할 수도 있지만 기계 학습과 별개로 순수하게 서비스 어플리케이션 개발만을 담당할 수도 있다. 그리고 서비스 사용자는 서비스 개발에 관여하지 않고 지능형 서비스를 이용한다. 제안하는 API 중 모델 생성 지원 API는 인텔리전트 데이터베이스에 저장된 학습대상 데이터와 자원을 이용하여 기계 학습을 수행하고 학습 모델을 생성하는데 필요한 API로 기계 학습 개발자가 주로 활용하게 된다. 그리고 서비스 개발 지원 API는 인텔리전트 데이터베이스에 저장된 데이터와 질의로 입력되는 데이터, 학습 모델을 사용하는데 필요한 API로 서비스를 개발할 때는 어플리케이션 개발자가, 서비스가 이루어질 때는 서비스 사용자가 이용한다. 인텔리전트 데이터베이스 사용자는 CLI(Command Line Interface) 또는 지능형 서비스 클라이언트를 통해 API를 사용한다.



그림 1. 제안하는 API 설계 목록

*교신저자 : yjs@chungbuk.ac.kr

2. 지능형 API 구현

본 논문에서 제안하는 API는 몽고 DB 데이터베이스를 기반으로 구현했다. 몽고 DB에는 Stored Procedure라는 개념이 존재하여 자바 스크립트를 데이터베이스 시스템 내에 저장할 수 있다. 그리고 저장된 스크립트는 몽고 DB 셸이나 명령어를 통해서 실행할 수 있다. 하지만 Stored Procedure를 이용하여 실행할 수 있는 스크립트는 자원을 많이 사용하지 않으며, 연산이 복잡하지 않은 작업만을 수행할 수 있다. 따라서 복잡한 기계 학습이나 케라스, 텐서플로우와 같은 기준에 만들어져있는 모델을 불러와서 사용하려는 경우 중간에 Web API 서버를 두어 동작시킨다. 그림 2는 지능형 API 동작 중에서 복잡한 연산을 수행하는 기계 학습 동작을 수행하기 위한 동작 과정을 보여준다. API 서버를 이용하여 동작시키는 경우, 몽고 DB에서는 해당 API를 호출하는 역할만을 수행하고, 실제 동작은 API 서버에서 수행하며 API 서버는 수행된 결과를 몽고 DB에 반환한다. 지능형 API 중에서 모델을 학습하는 동작을 수행하는 모델 생성 API의 대부분은 복잡한 연산과 많은 데이터에 대한 처리로 인해 많은 컴퓨팅 자원이 요구되기 때문에 API 서버를 통한 동작을 수행한다. 반대로 API 서버가 필요하지 않은 간단한 동작은 몽고 DB 자체적으로 저장된 Stored Procedure를 통해 수행한다. 지능형 API 중에서 모델 생성을 제외한 추론이나 데이터 처리 등을 API 서버를 통하지 않아도 동작한다.

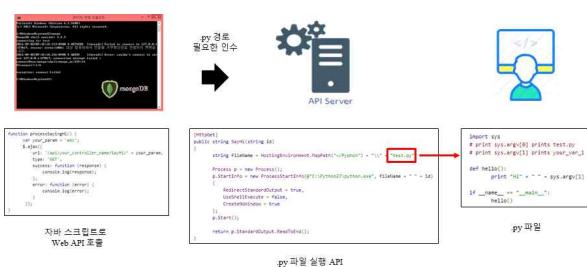


그림 2. 지능형 API 동작 과정 (복잡한 연산의 경우)

그림 3은 지능형 API 중에 모델 생성 API인 `modelCreation()`의 동작 과정을 보여준다. `modelCreation(model_info, data_info)`은 인수로 두 개의 값을 입력받으며, 해당 정보들은 JSON 형태로 기술되어 있다. 그림 4는 입력되는 `model_info`와 `data_info`를 보여준다. `model_info`는 모델 구성 정보를 의미한다. 모델 구성 정보란 레이어를 어떤 형태로 몇 계층을 구성할 것인지 등을 나타내는 정보이다. `data_info`는 학습 데이터 준비 정보를 의미한다. 학습 데이터 준비 정보란 MongoDB에 저장되어 있는 원본 데이터를 기계 학습에 입력할 수 있는 데이터 형태로 만드는 과정을 나타내는 정보이다. 해당 파일은 최초 모델을 생성할 때, 한번 입력을 수행하면 추후 성능에 따른 재학습 등을 수행할 때 다시 입력하지 않아도 되기 때문에 지속적인 서비스 제공을 하는 경우, 관리 비용이 감소된다.

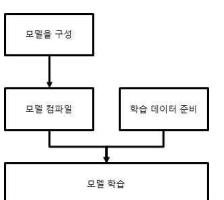


그림 3. 모델 생성 API

제안하는 API는 tensorflow.js에 따라 동작할 수 있도록 구현했다[3]. 그림 4에서 제시한 예제는 간단한 선형회귀를 수행하기 위한 모델 구성 및

학습대상 데이터 준비 과정이다. 해당 `modelCreation()`을 수행하면 그 결과로 생성된 모델이 뭉고 DB에 객체 형태로 저장된다. 저장된 모델은 모델 추론 API인 `modelInfer()`를 이용하여 실행할 수 있다. 그림 5는 뭉고 DB 셸에서 예제를 통해 생성된 모델을 실행하는 화면을 보여준다. `modelInfer(model_id, input)`는 인수로 두 개의 값을 입력받는다. `model_id`는 모델 생성 API를 통해 생성되어 저장된 모델의 id를 의미한다. `input`은 추론을 수행하는데 필요한 입력 값을 의미한다. 해당 예제를 통해 본 논문에서 제안하는 API를 통해 모델이 생성되고 추론을 하는 과정이 정상적으로 동작함을 확인했다. 그림 6은 복잡한 추론이 가능한지 확인하기 위하여 사람의 자세를 추정하는 기계 학습을 수행한 결과이다. 해당 기계 학습은 이미지 처리처럼 복잡한 처리가 요구되기 때문에 API 서버를 통해 수행했다.

```
C:\#Program Files\#MongoDB\Server\4.2\bin\mongo.exe
> db.loadServerScripts();
> modelInfer("5ea2773b43b81e5009951a8f", db.infertest.find())
Tensor[[8.3760147],]
```

그림 5. 모델 추론 API

그림 6. API 서버를 통한 모델 생성 및 추론

III. 결론

본 논문에서는 인텔리전트 데이터베이스에 기반을 둔 자율 기계 학습을 수행하기 위하여 지능형 API를 설계하고 구현했다. 제안하는 API는 모델 생성 지원 API와 서비스 개발 지원 API로 구성된다. 기계 학습 개발자와 서비스 개발자가 본 논문에서 제안하는 인텔리전트 데이터베이스 시스템에 기반을 둔 API를 활용하여 모델을 생성하거나 등록, 검색하여 서비스에 활용할 모델을 선정하고, 저장된 모델을 이용하여 추론을 수행하여 서비스에 활용하는 전체 과정을 수행할 수 있다. 추후 연구로는 tensorflow.js에 기반하여 모델을 생성하기 위한 레이어 구성 방법이 매우 다양하기 때문에 현재 구현되어 있는 기능에 추가하여 딥러닝을 지원하기 위한 더욱 다양한 레이어 구성 방법이 제안하는 API에서 동작할 수 있도록 할 예정이다.

ACKNOWLEDGMENT

이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보·컴퓨팅기술개발사업의 지원(No. NRF-2017M3C4A7069432), 중소벤처기업부 '산업전문인력역량강화사업'의 재원으로 한국산업기술진흥원(KIAT)의 지원(2019년 기업연계형연구개발인력양성사업, 과제번호 : S2755555), 그리고 2019년도 정부(과학기술정보통신부)의 재원으로 의 지원(No.B0101-15-0266, 실시간 대규모 영상 데이터 이해·예측을 위한 고성능 비주얼 디스커버리 플랫폼 개발)을 받아 수행된 연구임

참 고 문 헌

- ```
[1] https://www.tensorflow.org/
[2] https://cloud.google.com/automl?hl=ko/
[3] https://www.tensorflow.org/js?hl=ko
```