

## 2022-1학기 정보통신대학원 머신러닝2 실시간 비대면 기말고사 문제

시험시간: 2022.06.14(화) 20:15~21:05 (총50분)

### 유의사항

- 실시간 비대면 기말고사 안내파일을 다시 한번 더 읽어보고, 2022년 06월 14일(화) 오후 9시05분(시간엄수)까지 학번\_이름을 기입한 답안지 파일을 사이버캠퍼스 과제란을 통해 제출해 주길 바랍니다.
- 답안지 파일 형식은 pdf, ppt, 한글(.hwp), MSword(.doc 또는 .docx) 파일, 또는 수기로 작성하셔서 스캔이나 사진 찍어서 파일로 보내셔도 됩니다. 패드로 작성하셔서 pdf로 변환하여 보내셔도 됩니다. 어떤 형식이든 채점을 할 수 있도록 해 주시길 바랍니다.
- 문제는 총 20문제입니다.

다음의 설명들에 대해서 참(True) 또는 거짓(False)으로 답하십시오. 거짓(False)일 경우에는 그 이유를 설명하십시오. (여기에서의 참과 거짓의 판단은 본 강의의 강의자료 또는 강의동영상의 내용을 기준으로 판단하는 것을 말하며, 특별한 언급이 없는 한 모델(모형)은 신경망 기법 기반의 딥러닝 모델(모형)을 가정합니다). (문제당 각 5점)

1. 가중치 초기값은 대칭적 가중값으로 하여 중복성 문제가 발생하지 않게 하는 것이 좋다.
2. 최적의 매개변수를 찾아가는 방법 중 학습률은 그대로 두고 그래디언트를 갱신하는 방법으로는 AdaGrad, RMSProp, Adadelta 기법들이 있다.
3. 규제기법 중 가중치 벌칙은 규제 항에 가중치가 작아지면 더 강한 벌칙을 주도록 설계하여 가능하면 가중치가 작아지도록 설계하는 것이다.
4. 과적합(overfitting)을 방지하기 위해 입력층과 은닉층의 노드 중 일정비율을 임의로 선택하여 제거하는 방법을 드롭아웃(dropout) 규제기법이라 한다.
5. 데이터 셋을 훈련데이터, 검증데이터, 시험데이터 등 세 개로 나눌 때, 검증데이터는 하이퍼매개변수의 성능을 평가하기 위한 데이터로써, 일반적으로 시험데이터의 일부를 분리하여 사용한다.
6. 하이퍼매개변수 최적화를 위해 하이퍼 매개변수를 조합을 구현하는 방법으로는 격자탐색과 임의탐색이 있는데, 일반적으로 좋은 성능을 얻기 위해서는 일정한 간격으로 탐색하는 격자탐색이 더 우월하고 유리하다.

7. 적층 오토인코더 기법은 비지도학습을 수행하는 층별예비학습단계와 지도학습을 수행하는 MLP학습/미세조정 부분으로 나눌 수 있다.
8. 오토인코더 기법 중 결정론적 오토인코더 기법으로는 SAE, DAE, DBN 기법이 있고, 확률 오토인코드는 RBM, CAE 등의 기법이 있다.
9. 단어의 임베딩을 만들 때 백오브워즈 가정은 “어떤 단어가 같이 쓰였는가” 에 초점을 맞추고 있다.
10. 단어의 잠재의미분석(LSA)을 하기 위한 말뭉치의 통계량 정보가 들어있는 대상 행렬은 단어-문서행렬, TF-IDF 행렬, 단어-문맥행렬, 점별상호정보량행렬 등이 있다.
11. word2vec 기법 중 중앙의 단어(타겟)으로부터 주변의 여러 단어(맥락)을 추측하는 모델이 CBOW 모델이다.
12. word2vec 기법을 개선하기 위한 방법 중 부정적인 예를 샘플링하는 네거티브 샘플링 기법은 이중분류 문제를 다중분류 문제로 다루어 취급하며, 출력층의 손실함수를 소프트맥스 함수를 사용함으로써 어휘가 많아져도 계산량을 낮은 수준에서 일정하게 억제 할 수 있다.
13. 일반적으로 딥러닝모델에서 (i)다중분류문제인 경우, 출력층에서는 소프트맥스 함수를 활성화함수를 사용하고, 손실함수로는 교차엔트로피를 사용한다. (ii) 이진분류문제인 경우, 출력층에서의 활성화함수는 시그모이드 함수를 사용하고, 손실함수는 교차엔트로피를 사용한다.
14. RNN 학습시에 시간방향으로 펼친 신경망의 오차역전파법인 BPTT를 개선한 Truncated BPTT 기법을 제대로 구현하려면, 역전파의 연결은 그대로 유지하고 순전파의 연결만 끊어야한다.
15. RNN에서 Truncated BPTT의 미니배치 학습을 수행할 때는 데이터를 무작위로 선택해서 입력하여야하고, 미니배치별로 데이터를 제공하는 시작위치를 각 미니배치의 시작위치로 옮겨주어야한다.
16. RNN을 사용한 언어모델(RNNLM)의 예측성능을 평가하는 척도로 퍼플렉시티(perplexity, 혼란도)를 종종 사용하는데, 언어모델끼리의 비교에서 이 값이 작은 모델이 예측성능이 더 좋은 모델이라고 할 수 있다.
17. RNN 기법은 시계열 데이터에서 멀리 떨어진 장기 의존 관계를 기울기 소실문제 등의 이유로 잘 학습할 수 없다. 이를 해결하기 위하여 LSTM과 GRU 등 게이트가 추가된 RNN 계열 모형이 효과적이다.

18. LSTM 기법은 리셋 게이트와 업데이트 게이트로 총 두 가지 게이트가 있다.
19. 일반적으로 LSTM 기법의 게이트에는 전용가중치가 있다. 또한 활성화함수로는 주로 게이트에서는 소프트맥스함수가 사용되며, 실질적인 정보를 지니는 데이터에는 tanh 함수가 사용된다.
20. seq2seq 모델 구현에서 출발어 입력문을 인코더하는 Encoder에서 LSTM 계층은 오른쪽(시간방향)으로는 은닉상태와 셀을 출력하고 각 LSTM 계층 위쪽으로는 은닉상태만 출력한다. 이렇게 출력된 Encoder의 각 LSTM계층의 은닉상태는 이후 Decoder로 전달된다.

(문제 끝.)