



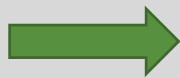
로또 번호 예측 PROJECT

AI_11_권진수

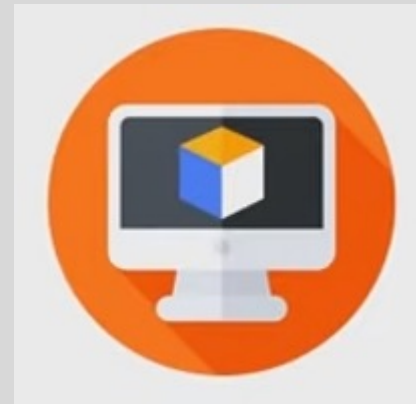
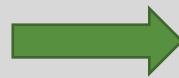
1. 로또 번호 예측 소개



1회~최근
당첨 번호 데이터



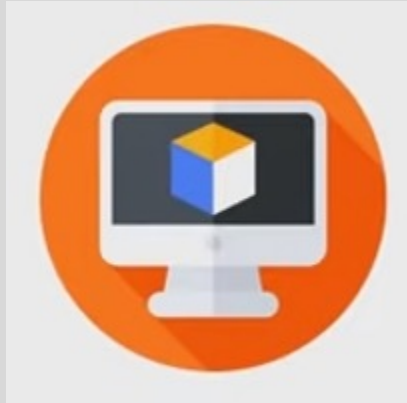
딥러닝 알고리즘
(LSTM)



학습 모델

1. 로또 번호 예측 소개

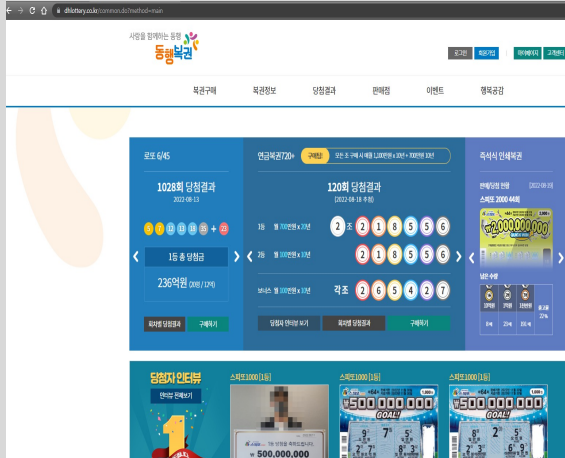
다음 회차
당첨 번호는 ?



[?, ?, ?, ?, ?, ?]

2. 데이터 수집

동행복권 : <https://dhlottery.co.kr/common.do?method=main>



순위	등위별 총 당첨금액	당첨개입 수	1개입당 당첨금액	당첨기준	비고
1등	23,624,702,640원	20	1,181,235,132원	당첨번호 6개 숫자일치	1등 자동18 수동2
2등	3,937,450,517원	91	43,268,687원	당첨번호 5개 숫자일치 +보너스 숫자일치	
3등	3,937,450,680원	2,810	1,401,228원	당첨번호 5개 숫자일치	
4등	7,138,900,000원	142,778	50,000원	당첨번호 4개 숫자일치	
5등	11,668,265,000원	2,333,653	5,000원	당첨번호 3개 숫자일치	

- 당첨금 지급기간: 지급개시일로부터 1년 (휴일인 경우 익명업일)

- 총판매금액: 100,613,537,000원

당첨번호조회

마수령 당첨금안내

1▼부터

1028▼까지

인쇄하기

엑셀다운로드

로또 공식 홈페이지 접속

당첨결과 -> 로또6/45 -> 회차별 당첨번호

1회부터 최근까지 엑셀 다운

3. 데이터 전처리

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	회차	1번	2번	3번	4번	5번	6번	보너스	1등 당첨금액	1등 당첨금액	1등 당첨금액	1등 당첨금액	1등 당첨금액
2	1	10	23	29	33	37	40	16	0	143934100	5140500	113400	10000
3	2	9	13	21	25	32	42	2	2002006800	94866800	1842000	100800	10000
4	3	11	16	19	21	27	31	30	2000000000	0	1174100	54900	10000
5	4	14	27	30	31	40	42	2	0	211191200	7282400	152100	10000
6	5	16	24	29	40	41	42	3	0	0	6033800	166500	10000
7	6	14	15	26	27	40	42	34	6574451700	196297600	4267300	87600	10000
8	7	2	9	16	25	26	40	42	0	48165000	1605500	65500	10000
9	8	8	19	25	34	37	39	9	0	131555000	2268100	110500	10000
10	9	2	4	16	17	36	39	14	0	769456500	8743800	260000	10000
11	10	9	25	30	33	41	44	6	6430437900	40813400	856400	27300	10000
12	11	1	7	36	37	41	42	14	4780152300	362132700	9307100	206800	10000
13	12	2	11	21	25	39	45	44	1348845700	99914400	1417500	70200	10000
14	13	22	23	25	37	38	42	26	0	433309300	5777400	175600	10000
15	14	2	6	12	31	33	40	15	9375048300	130363400	2568700	93900	10000
16	15	3	4	16	30	31	37	13	17014245000	177231700	5571100	144600	10000
17	16	6	7	24	37	38	40	33	4377146100	243174700	3385200	123500	10000
18	17	3	4	9	17	32	37	1	5349491200	297193900	2342100	86700	10000
19	18	3	12	13	19	32	35	29	0	175017300	3953700	136500	10000

각 회차 별 1~보너스 번호는 숫자이기 때문에 수치형 데이터로 착각하기 쉬운데 각 숫자는 크고 작음을 나타내는 것이 아닌 그냥 공을 나누어 각각 다른 숫자를 부여한 것이기 때문에 즉, 범주형 데이터 입니다.

3. 데이터 전처리

- 원핫인코딩

1	1000000...0000
2	0100000...0000
...	...
44	0000000...0010
45	0000000...0001



```
3 def numbers2ohbin(numbers):
4     ohbin = np.zeros(45)
5     for i in range(6):
6         ohbin[int(numbers[i])-1] = 1
7     return ohbin
8
9 def ohbin2numbers(ohbin):
10    numbers = []
11    for i in range(len(ohbin)):
12        if ohbin[i] == 1.0:
13            numbers.append(i+1)
14    return numbers
```

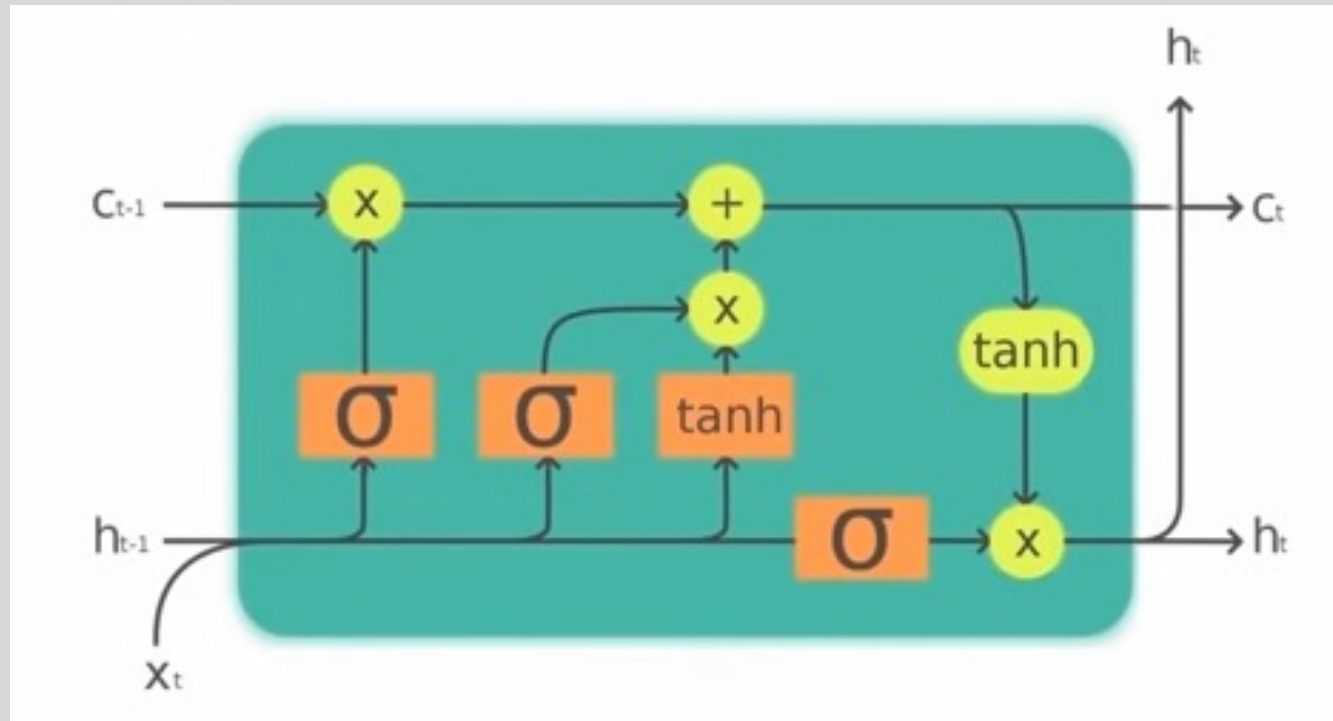
따라서 이런 범주형 데이터를 딥러닝에 편리하게 이용하기 위해 원핫인코딩을 실시합니다.

로또 번호는 총 45까지 이므로 전부 원핫 인코딩으로 바꿔줍니다.

4. 모델 구성 및 학습

- LSTM(RNN 계열)

시계열 데이터, 자연어 데이터 처리 등에 사용되는 딥러닝 모델



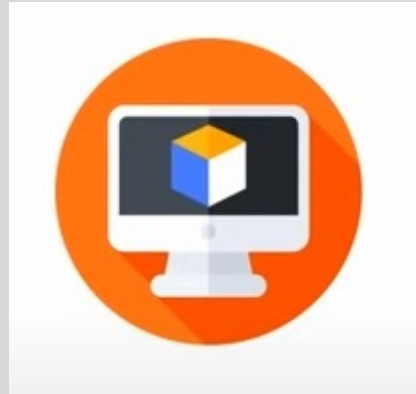
4. 모델 구성 및 학습

지도 학습

문제

이전 회차

당첨 번호



모델

정답

다음 회차

당첨 번호

5. 시연 영상

1. 몽고디비 – 데시보드
2. 학습을 통한 다음 회차 번호 예측

감사합니다.