



1 업무 이해

- 연구배경
- 현황파악
- 연구목적

 2
 데이터

 이해

- 정보수집
- 변수이해
- 정보통합

 3
 데이터

 준비

- 분할
- 변수설정

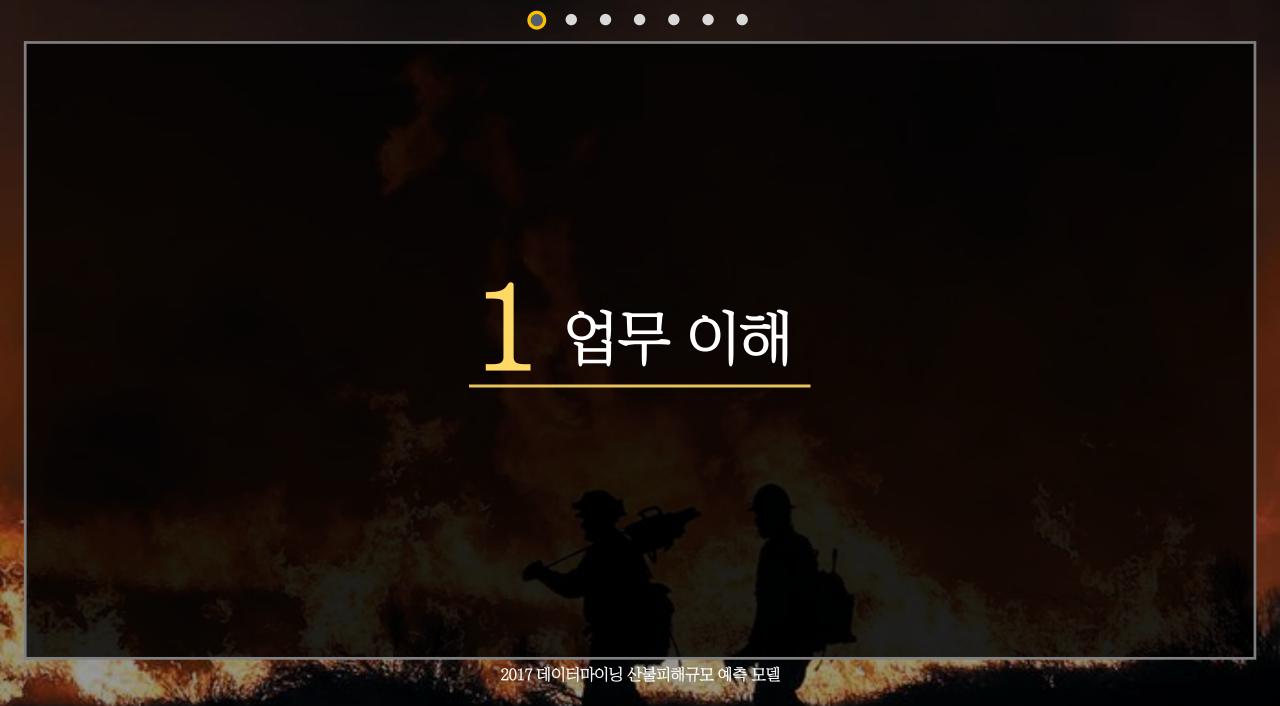
4 모델링

- 산불발생확률
- 산불피해규모

5 평가

6 결론

- 모델적용
- 향후과제



연구배경

영동지방 강풍에 산불 번져… 주민 2,500명 대피령



YTN NEWS https://www.youtube.com/watch?v=Ghp-qeK2aIA

연구배경

2017년 5월 6일~9일, 강원도 강릉지역 산불



산림 피해액 119 억 원



피해 주택 36 채



피해 면적 1,103 ha

산림청

산불 발생 추세



2017년도 면적 1,416 ha 건수 643 건

산림청

2016 산불피해액



대한민국 5대 재난성 산불

청양 예산 산불 (1996.04.23) 최대 풍속 15.1m/s

포항 울주 산불 (1996.04.23) 최대 풍속 15.1m/s 고성 산불 (1996.04.23) 최대 풍속 27m/s

양양 산불 (2005.04.04) 최대 풍속 32m/s

동해안 산불 (2005.04.04) 최대 풍속 23.7/s

산림청: 우리나라의 재난성 산불

산불과 기상요인의 관계

고온 건조한 강한 바람





대형 산불

삼성화재 방재연구소 : 이시영교수



기상요인에 따른 산불발생확률을 예측

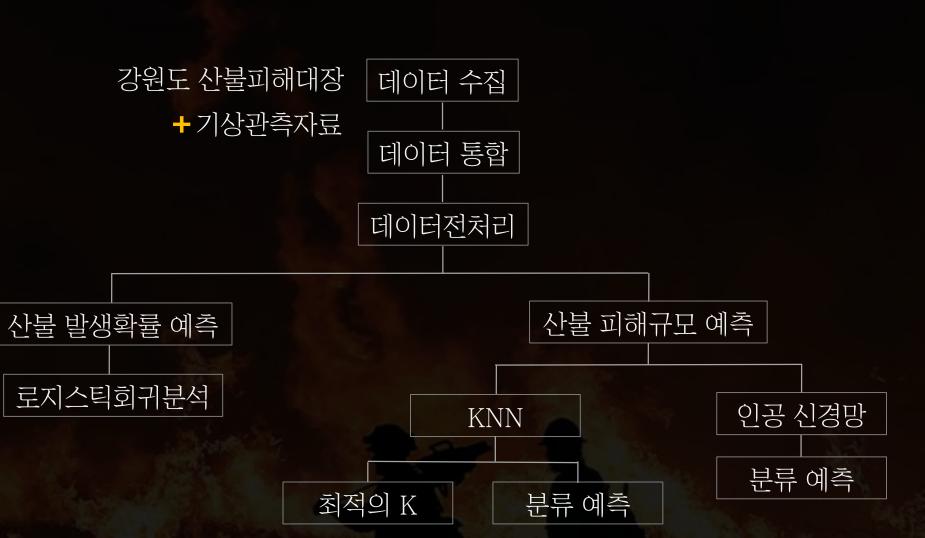
→ 산불위험도 제공으로 산불예방



기상요인에 따른 산불피해규모를 예측

→ 효과적인 단계별 진화대책 마련

연구 방법







데이터 수집 및 통합



기상자료개방포털

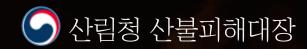




기점	시간	최고기온(°C)	일강수량(mm)	최대 풍속(m/s)	최소 상대습도(%)
속초(90)	2008-01-01	1.1		5.2	25
속초(90)	2008-01-02	6.1		5.9	19
속초(90)	2008-01-03	6.7		7	27
속초(90)	2008-01-04	7.5		3.8	25

연도	월	일	지역	지역	피해면적
2010	12	27	강원	영월	0.06
2010	12	27	강원	영월	0.5
2010	12	26	강원	강릉	0.6
2010	12	07	강원	인제	0.04





2008~2017 강원도 데이터 **498**개 19개 변수

년도별 산불	 불통계																		
발생일시_	발생일시_	발생일시_	발생일시_	발생일시_	진화종료시	진화종료시	진화종료시	진화종료시	발생장소_	발생장소_	발생장소_	발생장소_	발생장소_	발생장소_	발생원인_	발생원인_	발생원인	.피해면적_	합계
2017	12	01	12:55	금	2017	12	01	15:00	경북	경북	경주	내남	비지	산51	쓰	쓰레기소기		0.8	
2017	11	30	18:17	목	2017	11	30	19:10	경북	경북	포항	장기	죽정	산274	쓰	쓰레기소기		0.1	
2017	11	30	17:30	목	2017	11	30	20:30	전남	전남	영광	영광	도동	46-6	기	기타(직접(원인미상(0.1	
2017	11	30	17:20	목	2017	11	30	19:00	경북	경북	구미	고아	대망	산26	기	기타(직접(원인미상(0.6	
2017	11	30	15:26	목	2017	11	30	16:20	경남	경남	기장	남상	춘전	산76	기	기타(직접(주택 아궁	0.01	
2017	11	30	14:51	목	2017	11	30	16:48	경남	경남	창녕	남지	수개	산50	쓰	쓰레기소기		0.01	
2017	11	30	14:05	목	2017	11	30	14:58	경기	경기	수원	장안	상광교	산82-4	기	기타(직접(원인미상	0.13	
2017	11	30	11:14	목	2017	11	30	12:20	강원	강원	홍천	홍천	장전평	산82	쓰	쓰레기소기		0.05	
2017	11	30	10:38	목	2017	11	30	12:20	경남	경남	창녕	대지	효정	산109	<u>M</u>	쓰레기소격		0.01	
2017	11	29	13:10	수	2017	11	29	14:50	경남	경남	진주	명석	용산	산313	담	담뱃불실호		0.01	
2017	11	29	11:28	수	2017	11	29	16:00	전남	전남	여수	돌산	서덕	산9-1	쓰	기타(직접(
2017	11	27	15:00	월	2017	11	27	16:00	부산	부산	금정	금성		산2	쓰	기타(직접(낙엽소각	0.07	
2017	11	26	21:05	일	2017	11	26	22:25	경북	경북	경주	천북	갈곡	산58-1	성	성묘객실호		0.05	
2017	11	25	16:04	E	2017	11	25	17-20	겨나	겨나	사처	ОН	대성	YF22	71	기타/지저(의이미사	70 04	





강원도 지역별 일별 기상상태 2,217개 데이터 29개 변수

지점	일시	평균기온(°	최저기온('	최저기온	최고기온('	최고기온	일강수량(r	최대 순간	최대 순간	최대 순간	최대 풍속(최대 풍속	최대 풍속	평균 풍속	(풍정합(10	(평균 이슬	최소 상대 초	소 상대	·평균 상대	평균 증기
95	2015-01-01	-9.1	-11.6	751	-5.9	1459		7.7	320	1357	4.7	290	1321	2.1	1777	-20.4	30	1432	39.9	1.2
95	2015-01-02	-8.7	-13.6	743	-3	1441		7.2	290	1524	4.2	290	1541	1.2	1053	-17.7	27	1313	50.9	1.5
95	2015-01-03	-5.9	-16.1	744	0.7	1448		6.5	230	1435	3.8	200	1424	1.1	980	-12.9	34	1206	60.4	2.4
95	2015-01-04	-0.3	-5.6	2351	5.6	1409	0.2	4.4	230	336	2.9	230	344	0.8	688	-3	63	1	82.3	4.9
95	2015-01-05	-0.3	-7.5	534	6.9	1513	1.5	4.8	200	1614	2.9	200	1617	0.7	612	-4.2	44	1612	2 77	4.5
95	2015-01-06	-4.4	-10.9	2338	1.1	442	1.5	7.1	360	2114	4.4	360	2116	1.5	1333	-11.5	32	1554	59.4	2.9
95	2015-01-07	-9.5	-14.8	724	-2.8	1446		3.7	20	1424	1.8	20	155	0.7	573	-17.3	25	1330	56.5	1.6
95	2015-01-08	-9.8	-16.3	656	-2.2	1436		4.8	200	1514	3	230	1509	0.7	597	-16.2	29	1417	63.5	1.8
95	2015-01-09	-6.9	-14.7	729	1.8	1409		6	290	1433	3.6	270	1433	1	824	-13.8	19	1411	63.8	2.2
95	2015-01-10	-3.4	-13.6	752	2.9	1510		8.5	200	1523	4.6	230	1529	1.6	1406	-10.2	30	1425	62.6	2.9
95	2015-01-11	-1.2	-8.8	2350	2.3	450	0.1	7.9	320	551	4.3	290	552	2	1721	-8.9	33	1617	57	3.3



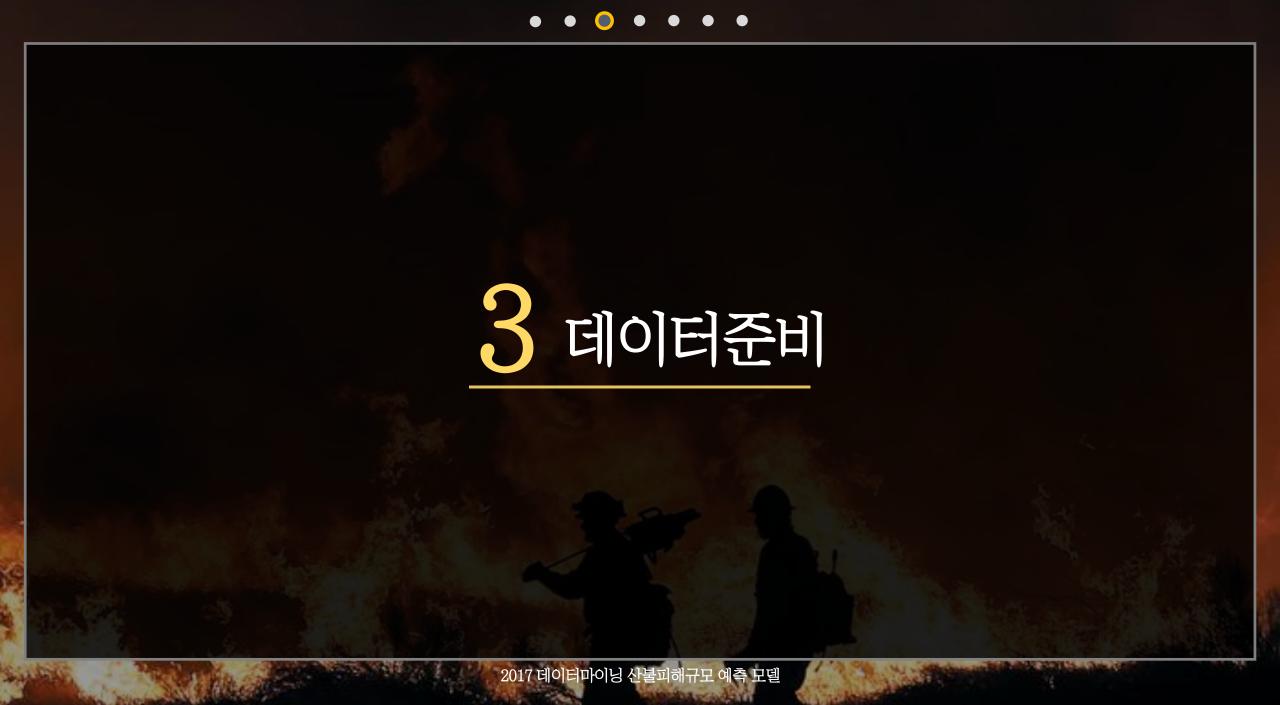




○ 산림청 산불피해대장 + ○ 기상자료개방포털 = 391개 데이터

연도	월	일	지역	지역	피해면적	최고기온	일강수량	최대풍속	최소상대습도
2010	12	27	강원	영월	0.06	-0.1	0	3.3	32
2010	12	27	강원	영월	0.5	-0.1	0	3.3	32
2010	12	26	강원	강릉	0.6	-0.7	0	10.6	15
2010	12	07	강원	인제	0.04	4	0	4.5	34
2010	12	04	강원	강릉	0.3	10.2	0	7.6	16
2010	12	03	강원	인제	0.3	9.8	3	6.2	24
2010	12	01	강원	강릉	0.01	19.5	0	4.9	21
2010	11	26	강원	춘천	0.1	6	0	1.7	13
2010	11	22	강원	춘천	0.3	11.6	2	5.1	21
2010	11	15	강원	원주	0.08	7.6	0	2.7	18
2010	11	12	강원	춘천	0.03	12.9	0.5	4.7	30
2010	10	24	강원	홍천	0.1	23.2	0	2.7	42
2010	06	24	강원	강릉	0.1	29.1	0	4.7	29
2010	06	10	강원	홍천	0.1	32.9	0	3.3	23

고온 건조한 강한 바람



데이터 변수 - 산불발생확률 예측

1)

25%, 4개 단계



Y

산불 발생여부

산불 발생 NO : 0 산불 발생 YES : 1

데이터 변수 - 산불발생확률 예측

1)

	fire	tempt	temp2 ²	temp3 ⁺	temp4	rain1 [®]	windf	wind2 ⁻	wind3 ⁻	wind4	rh1 [‡]	rh2 [‡]	rh3 [‡]	rh4 [‡]	
1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	
2	1	1	0	0	0	0	0	0	0	1	0	1	0	0	
3	1	1	0	0	0	0	1	0	0	0	0	1	0	0	
4	1	1	0	0	0	0	1	0	0	0	0	1	0	0	
5	1	0	0	1	0	0	1	0	0	0	0	0	0	1	
6	1	0	0	0	1	0	0	0	1	0	0	0	0	1	
7	1	0	0	0	1	0	0	1	0	0	0	1	0	0	
8	1	0	0	0	1	0	0	0	0	1	1	0	0	0	
9	1	0	0	0	1	0	1	0	0	0	0	0	0	1	



데이터 전처리 - 산불발생확률 예측

1)

훈련세트 60% 1130개

검증세트 40% 887개

데이터 변수 - 산불피해규모 예측

2)



산불 발생시,

- 최고기온
- 최대 풍속
- 최소 상대습도
- 일 강수량

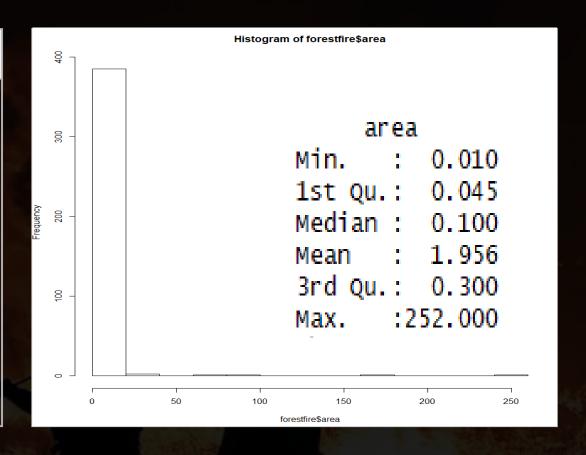


산불 피해면적

데이터 변수 - 산불피해규모 예측

2)

산불 피해면적



데이터 변수 - 산불피해규모 예측

2)

Y

산불 피해면적

단계	산불피해면적 (ha)
1 단계	0 ~ 0.1
2 단계	0.1 ~ 0.3
3 단계	0.3 ~ 0.7
4 단계	0.7 ~ 1
5 단계	1 ~ 2
6 단계	2 ~ 3
7 단계	3~



데이터 전처리 - 산불피해규모 예측

2)

```
> n=nrow(forestfire)
> n
[1] 391
> records=sample(1:n,round(n*0.7))
> train=forestfire[records,]
> test=forestfire[-records,]
> dim(train); dim(test)
[1] 274     7
[1] 117     7
```

훈련세트 70% 274개

검증세트 30% 117개



4 모델링

1) 산불발생확률 예측

로지스틱회귀

새로운 기상 데이터에 대한 산불발생확률 예측



확률 예측

로지스틱 회귀_

 $logit = -0.14829 - 0.66542 X_1 - 0.7162 X_2 + 2.24916 X_3 - 2.29235 X_4$

```
Coefficients: (1 not defined because of singularities)
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.14829
                        0.52447 -0.283 0.777372
temp11
            -0.66542
                        0.18611
                                 -3.575 0.000350 ***
            -0.12337
                        0.20038
temp21
                                 -0.616 0.538106
temp31
             0.05856
                        0.18898
                                  0.310 0.756654
temp41
                                     NA
                                               NA
rain11
            -0.71621
                        0.21162
                                 -3.384 0.000713 ***
wind11
             0.13468
                        0.40850
                                  0.330 0.741625
wind21
             0.03446
                        0.40650
                                  0.085 0.932439
wind31
            -0.02607
                        0.41101
                                 -0.063 0.949417
                                 0.618 0.536855
wind41
            0.26776
                        0.43357
rh11
             2.24916
                        0.46955
                                  4.790 1.67e-06 ***
rh21
            -0.01648
                        0.35774 -0.046 0.963249
rh31
            -0.50918
                        0.35828 -1.421 0.155269
                                 -6.372 1.86e-10 ***
rh41
            -2.29235
                        0.35974
```

```
> glmodel=glm(fire~.,data=forestfire, family = "binomial")
> summary(glmodel)

Call:
glm(formula = fire ~ ., family = "binomial", data = forestfire)

Deviance Residuals:
    Min     1Q     Median     3Q     Max
-2.1864   -0.4359   -0.3161   -0.2683     2.7819
```

```
\beta_0 = -0.14829

\beta_1 = -0.66542

\beta_2 = -0.71621

\beta_3 = +2.2491

\beta_4 = -2.29235

X_1 = \text{temp1}

X_2 = \text{rain1}

X_3 = \text{rh1}

X_4 = \text{rh4}
```

확률예측

로지스틱 회귀_

$$p = \frac{1}{1 + e^{-logit}} = 0.890987709$$

 $logit = -0.14829 - 0.66542 X_1 - 0.7162 X_2 + 2.24916 X_3 - 2.29235 X_4$

fire	temp1	temp2	temp3	temp4	rain1	wind1	wind2	wind3	wind4	rh1	rh2	rh3	rh4
1	0	0	0	1	0	0	0	0	1	1	0	0	0

산불발생확률 P=89.10%

4 모델링

2) 산불피해규모 예측

KNN

새로운 기상 데이터에 대한 산불피해규모 단계예측



KNN _ 변수 표준화

```
summary(forestfire)
                                                         rain
                      area
                                        temp
Min.
       : 1.0
                           0.010
                                   Min.
                                          :-5.00
                                                           : 0.0000
                Min.
                                                   Min.
1st Qu.: 98.5
                                   1st Qu.:11.85
                1st Qu.:
                          0.045
                                                   1st Qu.: 0.0000
Median :196.0
                Median :
                          0.100
                                   Median :20.70
                                                   Median: 0.0000
                                          :18.50
       :196.0
                      : 1.956
                                                           : 0.2072
Mean
                Mean
                                   Mean
                                   3rd Qu.:25.85
3rd Ou.:293.5
                3rd ou.: 0.300
                                                    3rd ou.: 0.0000
       :391.0
                        :252.000
                                          :35.10
                                                           :17.5000
Max.
                Max.
                                   Max.
                                                   Max.
     wind
                       rh
                                       step
       : 1.500
                        : 1.70
Min.
                 Min.
                                         :1.000
                                  Min.
1st Ou.: 3,600
                 1st Qu.:11.00
                                  1st Ou.:1.000
Median: 4.800
                 Median:17.00
                                  Median :2.000
      : 7.739
                         :18.69
                                         :2.143
Mean
                 Mean
                                  Mean
3rd Qu.: 6.550
                 3rd Qu.:24.00
                                  3rd Qu.:3.000
       :57,000
                 Max.
                         :71.00
                                         :7,000
Max.
                                  Max.
```

- > trctrl=trainControl(method="repeatedcv",number=10,repeats=3)
- > set.seed(3333)
- > knn_fit=train(step~.,data=train,method="knn",trControl=trctrl,
- + preProcess=c("center","scale"),tuneLength=10)

R, Carot 패키지 preprocess함수

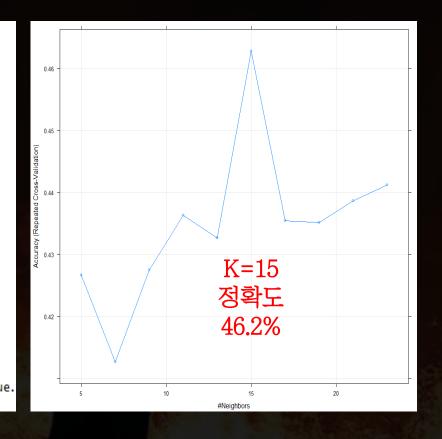
⇒표준화 작업

; 변수들의 범위크기가 다르기 때문



KNN _ 최적의 K 찾기

```
> knn_fit
k-Nearest Neighbors
274 samples
  6 predictor
 7 classes: '1', '2', '3', '4', '5', '6', '7'
Pre-processing: centered (6), scaled (6)
Resampling: Cross-Validated (10 fold, repeated 3 times)
Summary of sample sizes: 246, 246, 248, 249, 244, 245, ...
Resampling results across tuning parameters:
                Kappa
      Accuracy
     0.4266785 0.12641426
     0.4126735 0.09346813
     0.4274784 0.11269146
     0.4363141 0.12251427
  13 0.4326286 0.11332323
 15 0.4627823 0.15839728
 17 0.4354557 0.10962289
  19 0.4351303 0.10328652
  21 0.4386412 0.10735362
  23 0.4412062 0.10443385
Accuracy was used to select the optimal model using the largest value.
The final value used for the model was k = 15.
```





KNN _ 결과 분석

No Information Rate: 0.3932 P-Value [Acc > NIR]: 0.05478

Mcnemar's Test P-Value : NA

карра : 0.1761



KNN _

결과 분석

```
> confusionMatrix(test_pred,test$step)
Confusion Matrix and Statistics

Reference
Prediction 1 2 3 4 5 6 7
1 33 17 14 3 2 2 1
2 12 20 2 2 2 2 0 3
3 1 0 1 0 0 1 0
4 0 0 0 0 0 0 0 0
5 0 0 0 0 0 0 0
6 0 0 0 0 0 0 0
7 0 0 0 0 0 0 0

Overall Statistics
```

Accuracy: 0.4701 95% CI: (0.3772, 0.5645)

No Information Rate : 0.3932 P-Value [Acc > NIR] : 0.05478

Kappa: 0.1761 Mcnemar's Test P-Value: NA 정확도 47.01% P-Value 0.054

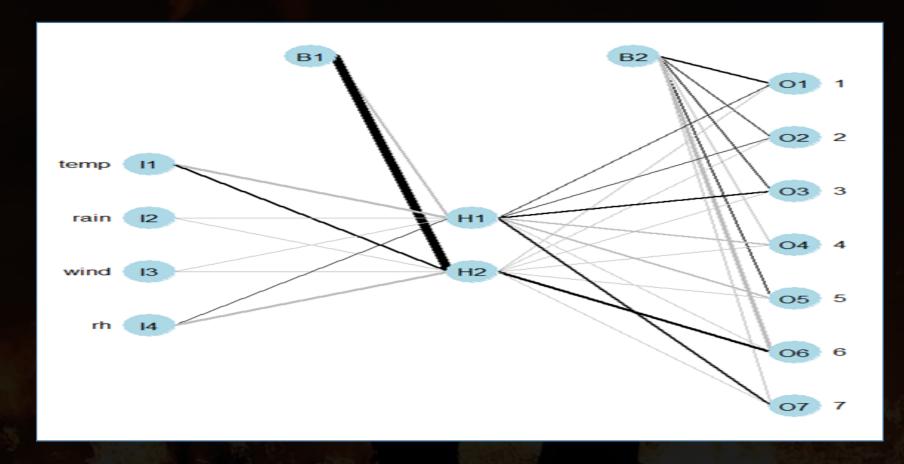
인공 신경망

새로운 기상 데이터에 대한 산불피해규모 단계예측

분 무 모 밀

인공신경망_

결과 분석





분류 모델

인공신경망_

결과 분석

```
> summary(modelnet)
a 4-2-7 network with 31 weights
options were - softmax modelling decay=5e-04
 b->h1 i1->h1 i2->h1 i3->h1 i4->h1
 -3.50 | -7.20 | -0.02 | -0.56
 b->h2 i1->h2 i2->h2 i3->h2 i4->h2
 16.16
        3.91 -0.07
                     -2.64
                            -6.69
 b->01 h1->01 h2->01
  2.86
        1.53 -1.91
 b->02 h1->02 h2->02
        0.80 - 0.75
 b->03 h1->03 h2->03
 1.88
         2.75 -1.68
 b->04 h1->04 h2->04
 -0.38 -3.20 -1.21
 b->o5 h1->o5 h2->o5
 0.47 -4.90 -0.66
 b->06 h1->06 h2->06
 -7.15 -0.35 6.32
 b->07 h1->07 h2->07
         3.36 -0.09
 -0.11
```

```
> erroverall=sum(cfmr)-sum(cfmr[diag.index])
> paste("오차율=",round(erroverall,digit=2),"%")
[1] "오차율= 59.59 %"
```

정확도 40.41%

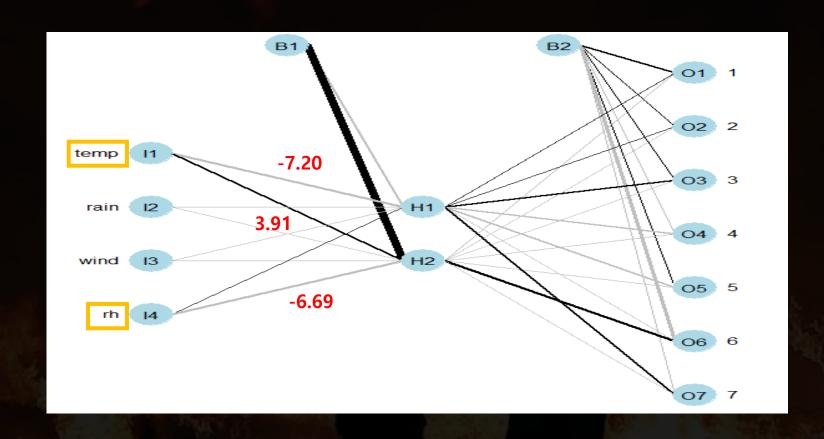
```
pred
actual 1 2 3
1 57 8 0
2 39 12 2
3 13 5 2
4 3 2 0
5 4 2 0
6 1 1 0
7 1 4 0
```



분 무 모델

인공신경망_

결과 분석



5 모델 평가



Evaluation _

인공신경망 모델보다 KNN모델이 더 유의미한 기법임을 확인.

KNN

정확도 47.01%

인공 신경망

정확도 40.41%



모델적용

산불발생확률 예측모델_

기상데이터를 기반으로 산불발생확률을 예측함으로써, 인근 지역에 산불 위험도를 제공하고 사전에 산불을 예방함.



기상데이터



산불발생확률예측



산불위험도제공

모델적용

산불피해규모 예측모델 _

실시간 기상데이터를 기반으로 중앙 산불정보시스템본부에서 여러 요인들을 분석하고, 산불피해규모를 예측함으로써 단계적으로 산불에 대응하도록 시스템 구축.



기상데이터



산불정보시스템본부



산불피해규모예측

한계

발화요인에 따라 불의 발생확률과 산불 피해규모가 다름을 고려하지 못했다.



산림청: 산불의 원인 및 영향

• • • • • • •

향후과제

- 1. 강원도 지역으로 국한되어 있는 현 모델을 전국 데이터기반의모델로 확장.
- 2. 기상요인 이외, 산림상태 및 발화요인 등을 분석하여 모델 정확도를 높임.

