

Calculating Crop Percentages and Predicting Price

Harshad B., Kuljot C. and Riya S.

PART 1 : Calculating Crop Percentages

The dataset used in this example was obtained from <https://aps.dac.gov.in>. The dataset contains information of area under a crop and its production from 2010-2018 for all districts in the state of Maharashtra. The dataset contains a wide variety of crops.

```
head(dataset)
```

```
##      Crop   District Year Season  Area Production Yield
## 1 Arhar/Tur Ahmednagar 2010 Kharif 17600      11300 0.64
## 2 Arhar/Tur Ahmednagar 2011 Kharif 12800       6400 0.50
## 3 Arhar/Tur Ahmednagar 2012 Kharif  9500       2200 0.23
## 4 Arhar/Tur Ahmednagar 2013 Kharif 11000       9200 0.84
## 5 Arhar/Tur Ahmednagar 2014 Kharif  9100       1800 0.20
## 6 Arhar/Tur Ahmednagar 2015 Kharif  9300       2200 0.24
```

Example

In this example, we will be looking at Ahmednagar district, and the season “Rabi” for the year 2010.

```
dist_in <- "Ahmednagar"
season_in <- "Rabi"
year_in <- 2010
```

We will subset the dataset according to our inputs.

We will write a function to calculate the percentages of crop under the particular selections.

Now, the results:

```
ans_df %>%
  arrange(Percent, desc = TRUE)
```

```
##      Crop      Percent
## 1      Sesamum 0.002323726
## 2      Linseed 0.007745753
## 3    Sunflower 0.061966027
## 4 Other Rabi pulses 0.340813149
## 5    Safflower 0.410524930
## 6      Maize 2.037133142
## 7      Gram 7.281008187
## 8      Wheat 19.279180189
## 9      Jowar 24.631495782
## 10 Small millets 45.947809114
```

PART 2 : Predicting Crop Prices

The dataset used in this example was obtained from <http://agmarknet.gov.in/>. The dataset contains information of the pricing of Bajra (Pearl Millets) in Nashik District and its various markets from 2010-2018 for all dates. The dataset contains minimum, maximum and modal prices of the Bajra crop in Nashik District for the given period.

Now, we will take a sneak peek at the data and while also clean the data (pre-processing) and change market names to numerical values for easier handling of data.

```
#taking a look at data before preprocessing
head(data)
```

```
##   i..Sr..No District.Name Market.Name Commodity Min.Price..Rs..Quintal.
## 1         1      Nashik    Chandvad    Bajra                1021
## 2         2      Nashik    Chandvad    Bajra                1001
## 3         3      Nashik    Chandvad    Bajra                1156
## 4         4      Nashik    Chandvad    Bajra                1011
## 5         5      Nashik    Chandvad    Bajra                1071
## 6         6      Nashik    Chandvad    Bajra                1129
##   Max.Price..Rs..Quintal. Modal.Price..Rs..Quintal. Price.Date
## 1                   1021                   1021   23022012
## 2                   1001                   1001   25012012
## 3                   1156                   1156    6012012
## 4                   1011                   1011   2122011
## 5                   1071                   1071  19112011
## 6                   1129                   1129  12112011
```

```
#convert dates from int to date
data$Price.Date <- as.Date(as.character(data$Price.Date), "%d%m%Y")
class(data$Price.Date)
```

```
## [1] "Date"
```

```
#remove NA dates
data$Price.Date[as.character(data$Price.Date) == 'NA'] <- NA
na_index <- which(is.na(data$Price.Date))
data_new <- data[-na_index,]
#change colnames
data_new <- data_new %>%
  rename(Sr.No = i..Sr..No,
         District = District.Name,
         Market = Market.Name,
         MinPrice = Min.Price..Rs..Quintal.,
         MaxPrice = Max.Price..Rs..Quintal.,
         ModalPrice = Modal.Price..Rs..Quintal.,
         Date = Price.Date) %>%
  mutate(Market = as.numeric(factor(Market))) %>%
  select(Market, MinPrice, MaxPrice, ModalPrice, Date) %>%
  arrange(Date)
#take a peek at the dataframe
head(data_new)
```

##	Market	MinPrice	MaxPrice	ModalPrice	Date
## 1	13	861	941	880	0011-02-11
## 2	8	898	1200	1045	0011-12-11
## 3	1	851	851	851	0011-12-21
## 4	8	900	1156	1021	0011-12-21
## 5	11	851	1296	1100	0011-12-21
## 6	13	848	1086	877	0011-12-21

The market codes are as follows:

```
markets <- data.frame(Market = unique(data$Market.Name), Factor = 1:length(unique(data$Market.Name)))
markets
```

##	Market	Factor
## 1	Chandvad	1
## 2	Devala	2
## 3	Dindori	3
## 4	Kalvan	4
## 5	Lasalgaon	5
## 6	Lasalgaon(Niphad)	6
## 7	Malegaon	7
## 8	Manmad	8
## 9	Nandgaon	9
## 10	Satana	10
## 11	Sinner	11
## 12	Suragana	12
## 13	Yeola	13

Example

In this example, we will be looking at Chandvad market.

```
target_market <- 1 #Chandvad
```

We will subset the dataset according to our inputs.

```
tm_data <- data_new %>%
  filter(Market == target_market)
```

We will train a neural network model with 15 hidden nodes and a threshold of 0.001.

```
net_price <- neuralnet(tm_actual[-1, 3] ~ Max_data + Min_data, data = trainingdata_price, err.fct = "sse")
```

And extract the modal pricing of the latest date available to us in the dataset.

```
testdata <- as.data.frame(tm_actual[nrow(tm_actual),])
net.results <- neuralnet::compute(net_price, testdata)
Predicted_Modal_Rate<- net.results$net.result
Modal_rate<- Predicted_Modal_Rate[1,1]
```

Now, the results

```
results <- data.frame(actual = testdata$Modal_data, prediction = net.results$net.result)
results
```

```
##      actual prediction
## 26    1021    1028.08
```