# Soft Sensor Design for Sulphur Recovery Unit

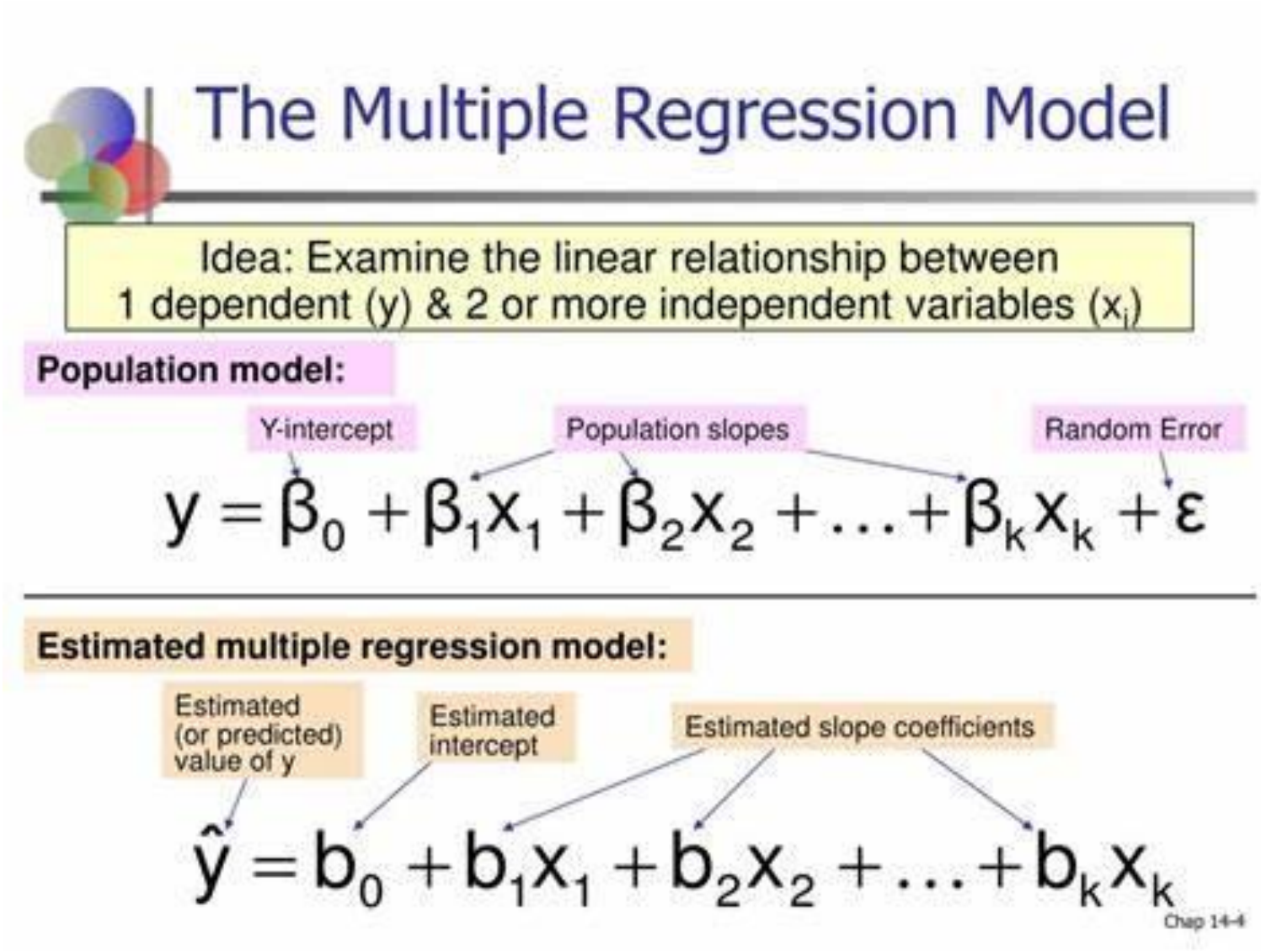
Under Prof. Jayaram Valluru

Abhishek Gonnade - 2022chb1037

# Algorithms applied

- Linear Regression
- Polynomial Regression
- KNN(K Nearest Neighbour) Regression
- Artificial Neural Network
- LSTM

# Multi - Linear Regression



# Multi - Linear Regression

| Predictive Performance | Values |
|---|---|
| Sum of Squared Errors (SSE) | 0.4136186523486100 6 |
| Mean Absolute Error(MAE) | 0.4719047082643700 3 |
| $R^2$ | 0.503788867504250 1 |
| Akaike Information Criterion (AIC) | -16761.41230124758 3 |
| Bayesian Information Criterion (BIC) | -16731.72007349175 7 |

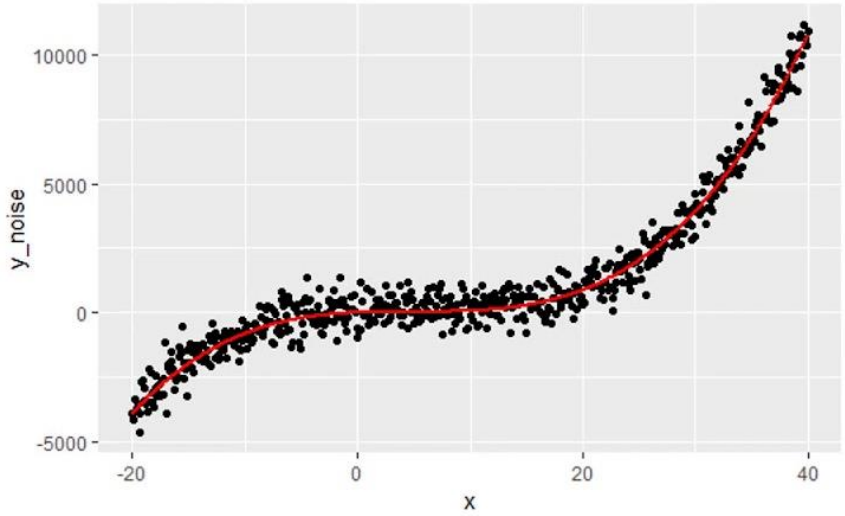
| $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|---|---|---|---|---|---|
| -0.010470 | 0.136875 | 0.15691 | 0.322204 | 0.036591 | 0.23094 |

# Polynomial Regression

- Used when relationship non-linear
- Can estimate a polynomial fit of higher-order
- User needs to define order

- $y = b + m_1 x^1 + m_2 x^2 + m_3 x^3$

- Sidenote: polynomial regression still type of linear regression ($x^2$ is just feature and $m_2$ a linear parameter)



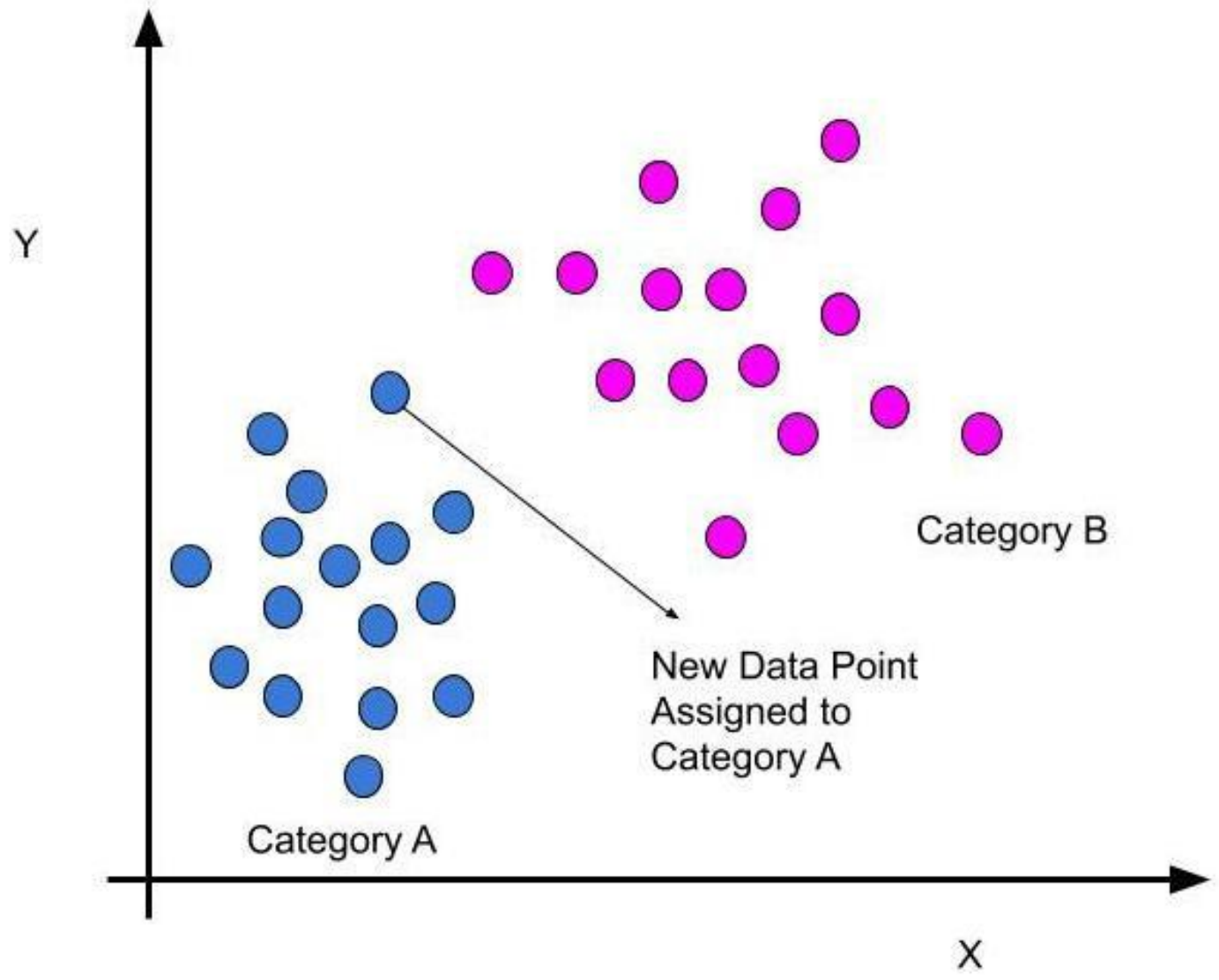
# Polynomial Regression

**For degree = 6**

| Degree | R2 score |
|---|---|
| 2 | 0.531 |
| 3 | 0.585 |
| 4 | 0.621 |
| 5 | 0.662 |
| 6 | 0.686 |

| Predictive Performance | Values |
|---|---|
| Sum of Squared Errors (SSE) | 0.26149008571890997 |
| Mean Absolute Error(MAE) | 0.3801269169579971 |
| $R^2$ | 0.6862948737098291 |
| Akaike Information Criterion (AIC) | -18046.72248558417 |
| Bayesian Information Criterion (BIC) | -18017.030257828344 |

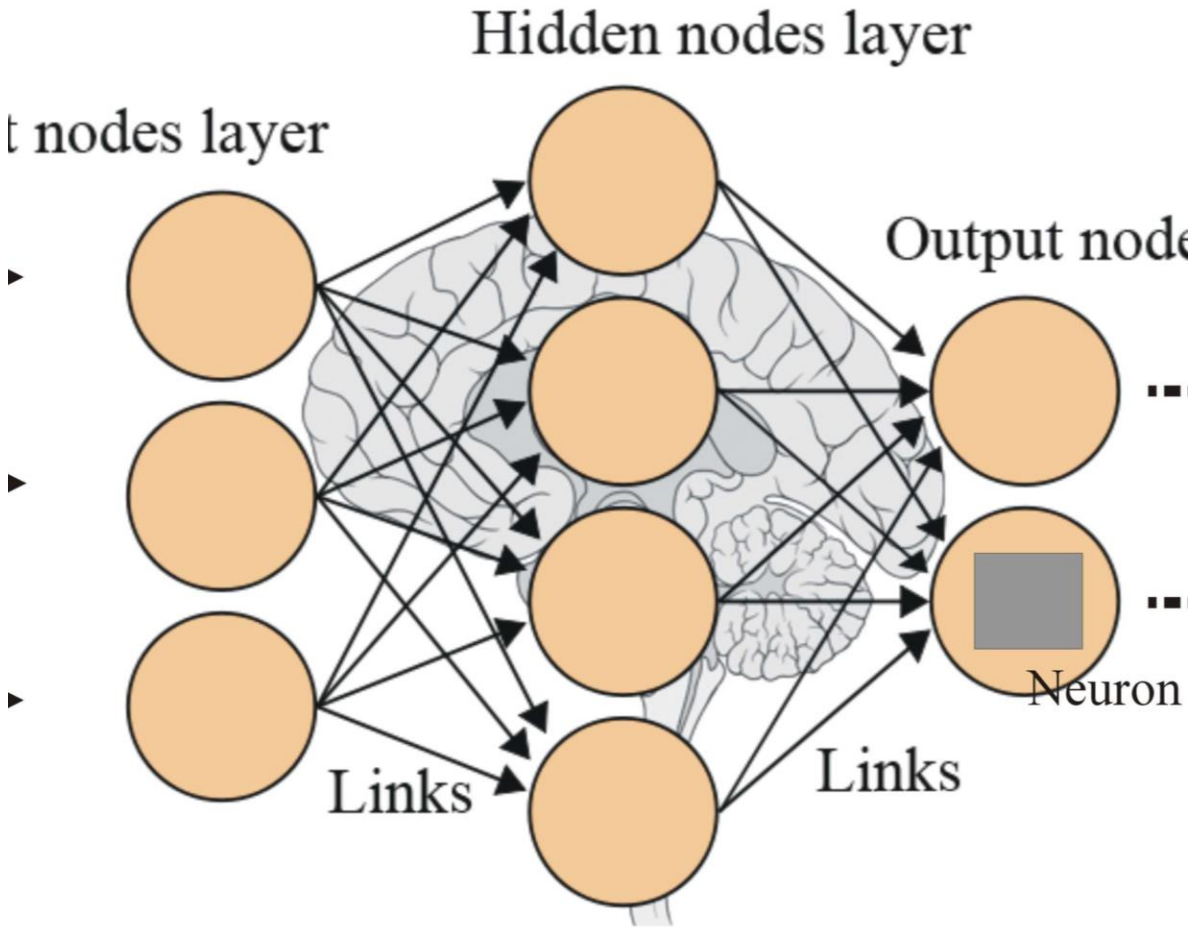
# KNN(k – Nearest Neighbour)



# KNN (k - Nearest Neighbour)

| Predictive Performance | Values |
|---|---|
| Sum of Squared Errors (SSE) | 0.253710086927959 1 |
| Mean Absolute Error(MAE) | 0.3330039616920311 7 |
| R2 | 0.695628403493731 5 |
| Akaike Information Criterion (AIC) | -18131.38470283076 5 |
| Bayesian Information Criterion (BIC) | -18101.69247507494 |

# ANN(Artificial Neural Network)

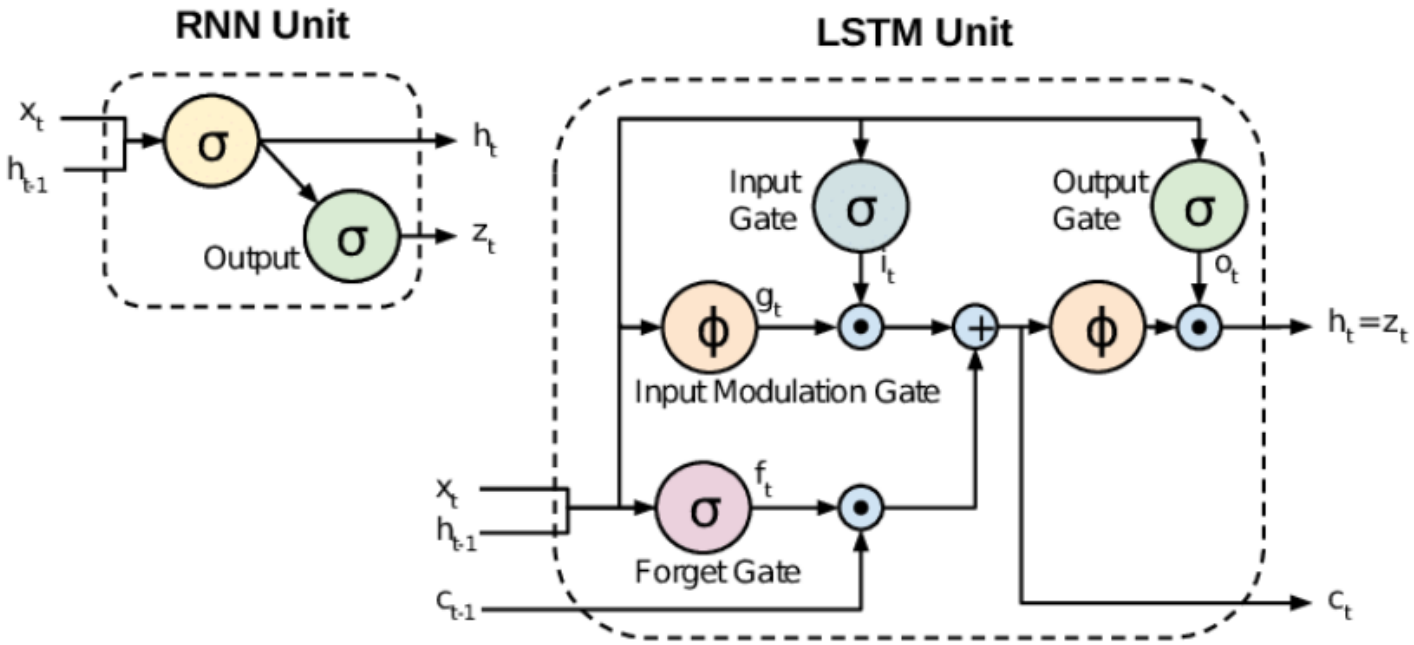


# ANN(Artificial Neural Network)

Number of Epocs =300

| Predictive Performance | Values |
| --- | --- |
| Sum of Squared Errors (SSE) | 0.20884828958355764 |
| Mean Absolute Error(MAE) | 0.3279026738079866 |
| $R^2$ | 0.749448324669090 1 |
| Akaike Information Criterion (AIC) | -18676.804002033543 |
| Bayesian Information Criterion (BIC) | -18647.111774277717 |

# LSTM


RNN Unit
LSTM Unit
Output
Input Gate
Output Gate
Input Modulation Gate
Forget Gate

# LSTM

Number of Epocs = 1000

| Predictive Performance | Values |
|---|---|
| Sum of Squared Errors (SSE) | 0.17731212350159817 |
| Mean Absolute Error(MAE) | 0.3016487145636578 |
| $R^2$ | 0.7872817168462738 |
| Akaike Information Criterion (AIC) | -19135.645325211815 |
| Bayesian Information Criterion (BIC) | -19105.95309745599 |

# Conclusion

| Model | $R^2$ score |
| --- | --- |
| Multi-Linear Regression | 0.503 |
| Polynomial Regression | 0.686 |
| KNN | 0.696 |
| ANN | 0.749 |
| LSTM | 0.787 |