E0 259
Data Analytics
Assignment 1: D/L Method

Mayank Sati
SR No. : 19264

1 Libraries Imported:

- **import numpy as np**
- **import pandas as pd**
- **from matplotlib import pyplot** (for graph)
- **from scipy import optimize as op** (for minimize function)
- **import math** (for exp)

2 Implementation Summary

2.1 Data Cleaning

The necessary columns needed were:

- Over
- Innings.Total.Runs
- Total.Runs
- Wickets.in.Hand

Also, since parameter 'u' is Overs to Go instead of Overs Bowled, I modified the column 'Over' to (50-'Over').

2.2 Optimization

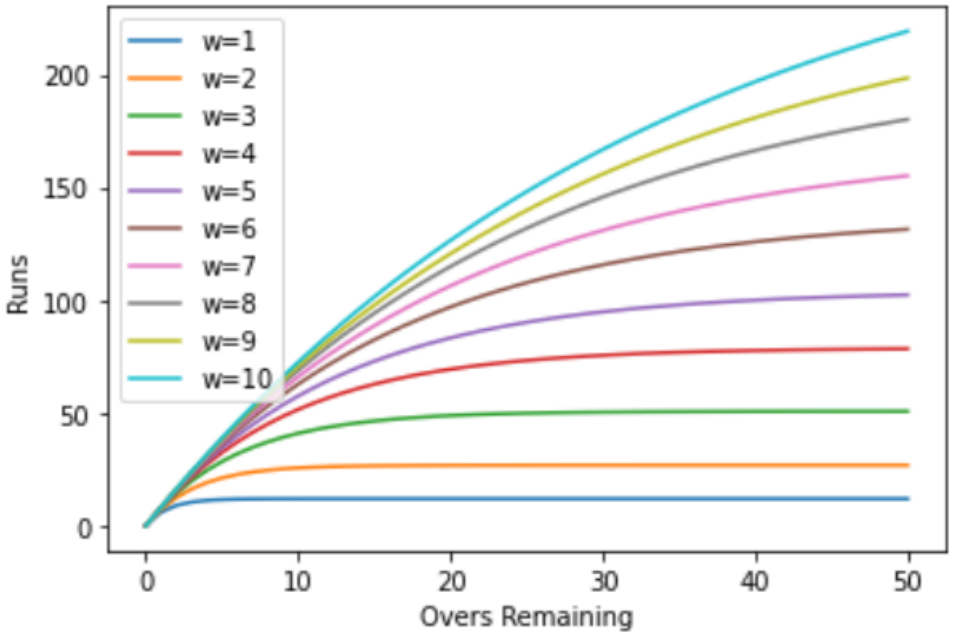
The optimization function is **Squared Error Loss** i.e the total squared error across all data points for the overs and wickets, where the predicted value for u,w is given as:

$$Z(u, w) = Z_0(w)[1 - e^{-Lu/Z_0(w)}]$$

I used scipy.optimize.minimize function for non linear regression.

2.3 Reports and Graphs

Following is the plot generated.



The 11 Parameters Reported are:

Wickets in Hand	Z_0
1	11.6664
2	26.8078
3	50.6184
4	78.5794
5	103.9465
6	137.6539
7	168.8422
8	207.5715
9	239.1373
10	284.2157

L	10.8822

And the Normalized Squared Error (after finding optimal parameters) across all overs and wickets is gven as:

Normalized Squared Error	381.2825