**Team Update re Project Proposal**

**I have copied this from the Project Proposal to highlight what we already have team. Note we satisfy all of the requirements for the Final Project and then some!**

**The good news is that we can complete this in a timely manner and focus on learning the machine learning / regression stuff so that we are all singing from the same hymn book at our presentation.**

1. **Project Brief Description**:

The project uses the machine learning approach to create a model for analysing and forecasting the Covid-19 Pandemic.

The project:

* performs Extract, Transform and Load to extract Covid-19 data from the John Hopkins University time series data sets which are published daily,

I ran this but we need to double check the data extracted. I’m not sure that it is all there. Will discuss when we meet. We need to tidy up the “create-covid-db” code.

* stores the data sets in a PostGreSQL data base,

We need to revalidate the DB Schema, the Set Up Code, and the Load.

We also could develop some SQL queries to go with the DB

* assesses various regression model techniques to find the **optimal regression model** for analysing and forecasting Covid-19 Confirmed Cases, Active Cases, Recovered Cases and Deaths,

I am well down the track with this. Will provide an update and demonstrate tonight. BTW the optimal regression model is Polynomial Regression with 5th degree Polynomial in X. Sounds impressive doesn’t it.

I have completed the evaluation of the Linear Regression model using the Covid data and the “Lasso”, “Ridge” and “ElasticNet” models.

I have yet to complete the evaluation of the Polynomial Regression models and need help from Uday / Saheed on this.

Note we are all going to have to go through the optimal regression model until we are completely comfortable with it and can answer any questions at our presentation.

* uses time series analysis and the **optimal regression model** identified above to produce forecasts of Covid-19 Confirmed Cases, Active Cases, Recovered Cases and Deaths,

I have got the basic optimal regression model working but need some help from Uday / Saheed to ensure it is correct.

* uses **Tableau** to read the forecasts produced above and **apply its exponential smoothing for forecasting and plotting visualisations** of Covid-19 Confirmed Cases, Active Cases, Recovered Cases and Deaths,

I have a sample Tableau forecast to show you tonight. It needs fine tuning when we get the data sorted.

* compares the results of the **optimal regression model** to the results produced by **Tableau**.

Yet to be done.