Niall larkin

Niall.Larkin@city.ac.uk

INM433 Visual Analytics Individual coursework submission:  
Visual Analytics analysis predicting Iron Ore quality in Iron Ore flotation processes

By Niall Larkin

# Motivation, data and research questions:

## Motivation for study and domain specific research questions for investigation:

The motivation for this study is to assess the quality of Iron ore feed output from an Iron Mining Silica removal froth flotation unit operation based on the quality of its feedstock and unit operation parameters. As a result of this there a number of analytical questions that come from this objective:

* What data from the manufacturing sensors are genuine feedback on the process?
* Is it possible to predict % Silica Concentrate every minute based on the process parameters and feed stream stock?
* How many steps (hours) ahead can we predict % Silica in Concentrate?
* Is it possible to predict % Silica in Concentrate without using % Iron Concentrate column

This analysis is pertinent from a business standpoint as by changing optimising the processing with regards to understanding Silica concentration provides values in 2 ways:

1. Having the ability to reduced down the cost of finish product sampling reduces of the overall manufacturing process and decreases batch release time for each lot of ore.
2. Having a better characterisation of how different processing parameters influence the final silica output leads to the opportunity that silica removal can be further optimised improving downstream processing of iron ore by reducing the energy requirements during pig Iron formation[XX]

## Data suitability:

The data set utilised for this analysis is data directly from the manufacturing shop floor of mining unit operation. As a result it is suitable dataset to utilise for this analysis. One item to be conscious off the time series data output stream results for iron and silica are taken on an hourly basis and are meant to represent the subsequent hour of material as a result.

## Data Transformations

There were a number of data transformations that where required prior to performing any analysis on this dataset.

1. Removal of all commas with decimal places in the dataset
2. Change date timestamps as the time measurement for the process parameters where taken every 20 seconds and therefore need to be changed to reflect this for any time series analysis.
3. Creating two new parameters
   1. % Iron feed enrichment: This was the subtraction of the final iron feed concentration from the initial feed
   2. % silicon feed removal: This was the subtraction of the final silica feed concentration from the initial feed

# Tasks and approach

## What manufacturing instrumentation signals are genuine processing signals

From a manufacturing standpoint typical issues there a myriad of issues encountered with regards to poor or noisy feedback from instrumentation due to manufacturing conditions. As a result two activities need to be performed to address this:

1. Data quality assessments: For poor signal feedback for the process instrumentation or analytical errors are commonplace in a manufacturing setting. This is typically easy to identify by a lack of variance within the time series trend where it is physically impossible to have such phenomenon. In order to determine where this is present all time series trends will be plotted where static no variance in the trends will be identified and removed from the data set for analysis.
2. For noisy time series data with PACF and ACF trends need to be generated to determine the time frame for the smoothing variable. The type of smoothing that will be used due to the fact that all the processing signals are digital in nature will be the savtivzky golay filter which is well adapted for data smoothing for electrical signals. The smoothing process for this will be treatable however with a visualisation of the trend of the dataset being generated and the associated order of the filter being changed until an appropriate fit is observed. From this then smoothing moving average smoothing can be performed to remove the associated noise. If no seasonality the distribution will be assessed for white noise by reviewing a residual plot of the time series. For froth flotation this typically the process control parameter pulp flow or tank level which will be determined by time series analysis for stationarity[Refernece].

Is it possible to predict % Silica Concentrate every minute based on the process parameters and feed stream stock?

Once the level of signal noise has been smoothed and ineffective signals have been removed a effective visual representations of the different parameters within the data can be visualised. Due to the process flow present of this process(See figure XX) the analysis of this process can be broken down into two analytical question subtasks:

1. What feed stock parameters influence the final % silicon concentration
2. What feedstock dosing parameters influence the final % silicon concentration.
3. What manufacturing process parameters influence the final % silicon concentration.

This can be assessed visually using regress pairplots and cluster analysis if clusters are present as opposed to linear or non-linear correlations.

## Manufacturing processing parameters

## Output feed steam characterisation.

# Analytical steps:

## Data quality assessment:

Initial review of the datasets revealed a number of areas where data quality was called in question specifically with regards to the quality control measurements taken from the finished product(Please circled area highlighted in Figure 1 to Figure 2 below). The typical distributions of the finished product testing excluding these results can be seen in Figure 3. This is typical that analytical tests of mineral samples take on a normal distributions in part due to the variability of the minerals and the analytical test itself[XX]. The fact that consistent outliers are present and relay a consistent stream in this distributions for in excess of 56.25 hours for both silica and iron assay testing at the same time is not physically possible From this it can be observed these results are an outlier and the fact that both results are consistent is physically impossible as all analytical tests

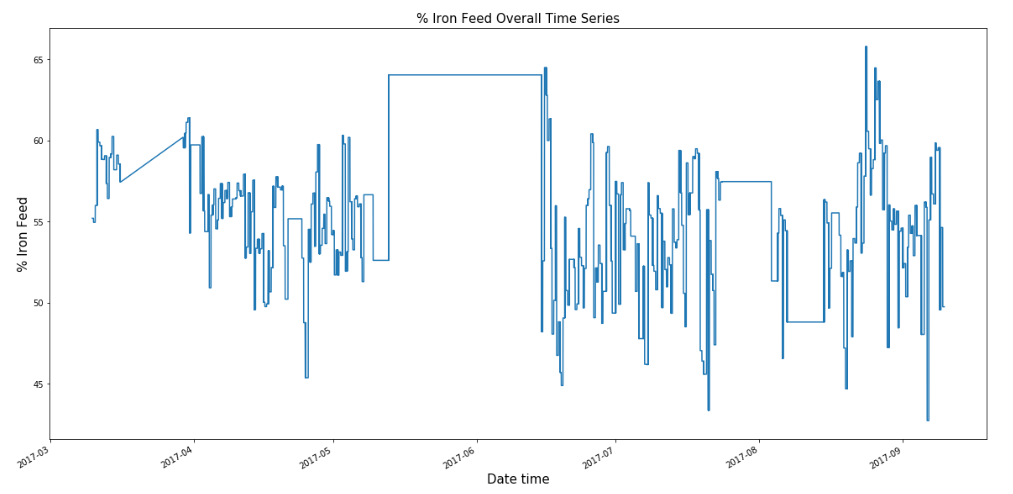


Figure Points of static feedback from quality control measurements of% w/w Iron feed into froth flotation tank

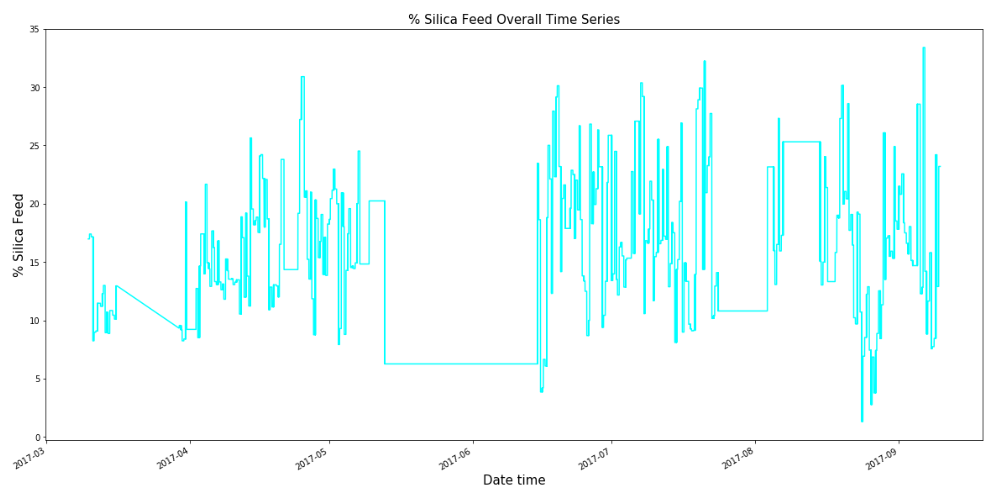


Figure Points of static feedback from quality control measurements of% w/w Iron feed into froth flotation tank

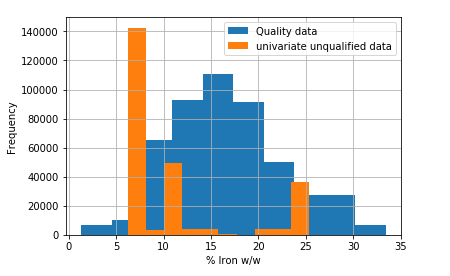


Figure Distribution of %Iron w/w assay values of feedstock quality distributed data and unqualified univariate data are ploted to show that the data is primarily outliers