A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Jaiber Camacho-Olarte
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Jaiber Camacho-Olarte 1,*, Julián Esteban Salomón Torres 1, Daniel A. Garavito Jimenez 1, Jersson X. Leon Medina 1, Ricardo C. Gomez Vargas 1, Diego A. Velandia Cardenas 1, Camilo Gutierrez-Osorio 1, Bernardo Rueda 2, Whilmar Vargas 2, Diego Alexander Tibaduiza Burgos 1, Cesar Augusto Pedraza Bonilla 1, Jorge Sofrony Esmeral 1, Felipe Restrepo-Calle 1

1 Universidad Nacional de Colombia, Bogotá, Colombia.
2 Cerro Matoso S.A., Montelíbano, Colombia.
* Correspondence: dtibaduizab@unal.edu.co

7th International Electronic Conference on Sensors and Applications
15 – 30 November 2020
INTRODUCTION

Cerro Matoso S.A. Smelting Company
South32 - Cerro Matoso

- Mining, Production and export of nickel
- Operating since 1982
- Open pit mine

1880 Employees
40,000 Tonnes Ni
Average production last 3 years

1st Producer FeNi LATAM
Total revenue FY20 500 US$M
10% South32
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

South32 - Cerro Matoso

Ferronickel Production Process

Mining and Homogenization  Drying and Sorting  Calcination  Smelting  Refining  Finished Product
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Adapted from Forero Caro, J.C., 2019
LITERATURE BACKGROUND.
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Different Approaches for Data Cleaning Process.

- Statistical Analysis
- Algorithms for time series
- Integrity Constraints
- Neuronal Networks
- Machine Learning Models

What to detect?

How to detect?

Where to detect?
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

DATASET

- Hundreds of measurements per minute.
- CSV file with 1180 variables and 175297 rows.
- Size of file: 2.8GB

Experts

15-minutes Frequency
4 years

12
DATA CLEANING PROCESS.
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Workflow for Data Cleaning Process.

- Remove duplicates
- Empty and null values
- Unique values
- Strings

<table>
<thead>
<tr>
<th>Time</th>
<th>Raw data</th>
<th>Processed</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30</td>
<td>ON</td>
<td>1</td>
</tr>
<tr>
<td>12:45</td>
<td>ON</td>
<td>1</td>
</tr>
<tr>
<td>13:00</td>
<td>Failure</td>
<td>null</td>
</tr>
<tr>
<td>13:15</td>
<td>OFF</td>
<td>0</td>
</tr>
</tbody>
</table>
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Workflow for Data Cleaning Process.

- **Negative temperatures**
  - TAG value vs. time
  - Diagram showing fluctuations

- **High variance**
  - TAG value vs. time
  - Diagram showing high peaks and valleys

- **Zero variance**
  - TAG value vs. time
  - Diagram showing flat or constant trend
05

RESULTS
Amount Detected Variables in Data Cleaning Process.

- **Remove duplicates**
  - 80 duplicated TAGs
  - 5 TAGs were eliminated

- **Empty and null values**
  - 2 empty TAGs
  - 74 TAGs were eliminated

- **Unique values**
  - 60 TAGs were found

- **Strings**
  - 2 empty TAGs

- **Negative temperatures**
  - 74 TAGs were eliminated

- **High variance**
  - 97 TAGs were found

- **Zero variance**
  - 22 TAGs were found
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

Principal Results

314 Eliminated Variables

28% Dirty Data

1.1 GB Size of file reduced
CONCLUSIONS
CONCLUSIONS

- Workflow for a data cleaning process.
- The expert judgment is very important.
- Data quality was improved.
REFERENCES


Questions?

jfcamachoo@unal.edu.co
Universidad Nacional de Colombia
2020
A Data Cleaning Approach for a Structural Health Monitoring System in a 75 MW Electric Arc Ferronickel Furnace

South32 - Cerro Matoso