SYMPOSIUM ON FERROUS METALLURGY 2024

Rationalisation of steel grades and specifications

Sadegh Jalalian [a], Hamid Assadi [a], Isaac Chang [a], Zhongyun Fan [a]

[a] Brunel University London, BCAST

Steel, Material Science, Artificial intelligence, Machine learning, Classification …

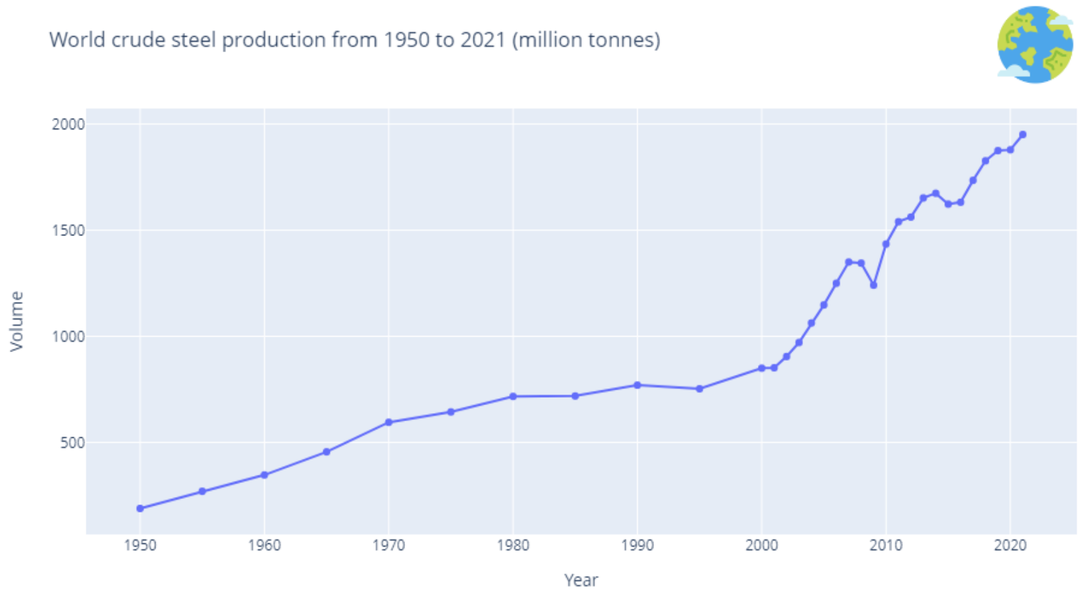
Brunel University London | UKRI Interdisciplinary Centre for Circular Metals | UK Research and Innovation | NICER PROGRAMME

Steel Background

- Steel has been used since the 19th century.
- Different steel grades are used to distinguish the different types of steel based on their unique properties.
- Steels are conventionally divided into four groups, including alloy steel, carbon steel, stainless steel, and tools steel.

Aim of the project

This project will aim to provide an analytical basis for reclassification of steels into a significantly smaller number of grades, to improve recyclability without compromising performance and production costs.



Source: World Steel Association

In the existing steel groups, some of the steel grade's performances are slightly the same in the production phase and differ in composition and processing conditions.

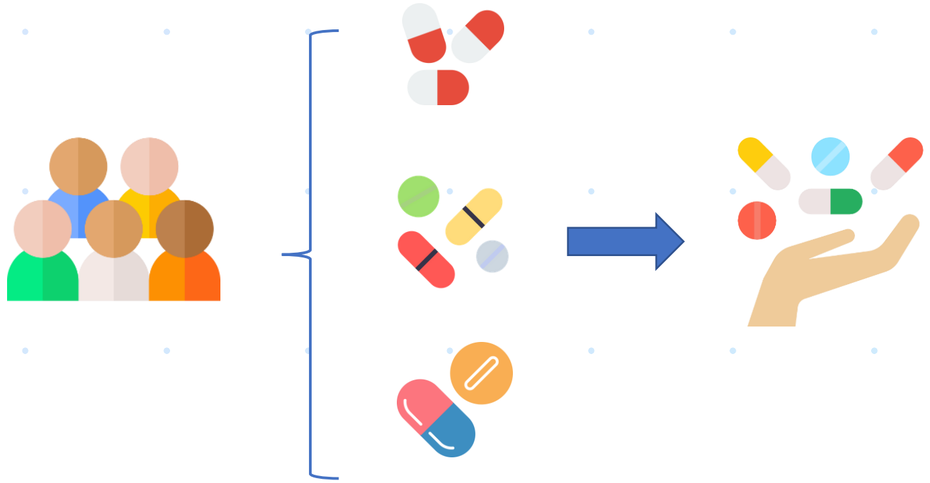
Raise Production Costs

Cause Environmental Damage

Make recycling end-of-life products challenging

Insight Into The Classification Context

Classification is defined as the process of recognition, understanding, and grouping of objects and ideas into preset categories a.k.a “sub-populations.”

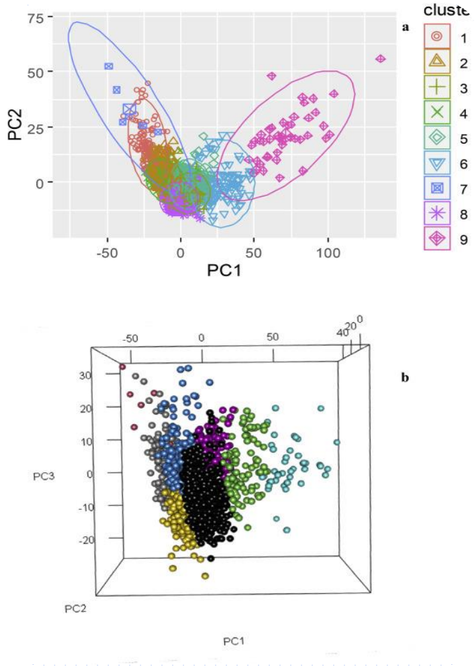


Description

Drug discovery is the process of making drugs for the targeted disease (analogous to engineering application of steel), with minimal side effects on the patient’s body (analogous to negative environmental impact). Lead compound identification (analogous to alloy composition) is one of the essential steps for finding the relevant drug for a particular disease in the drug discovery process.

Drug Discovery

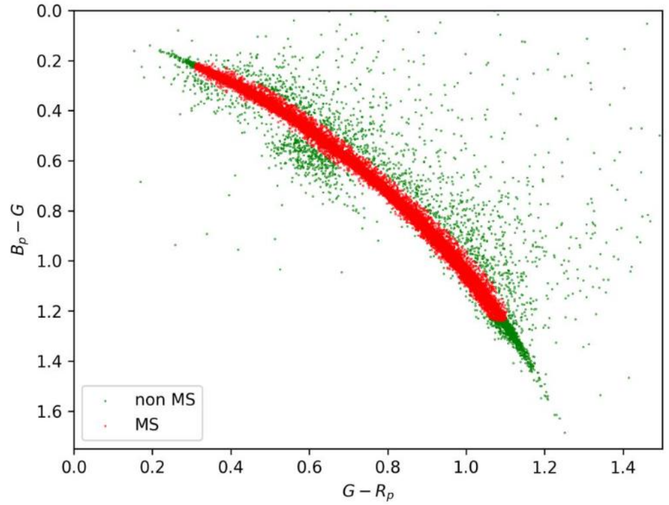
Molecular descriptor analysis of approved drugs using unsupervised learning for drug repurposing.



The purpose of this study was to analyse approved drug information that contained molecular and biological activity for drug repurposing.

Astronomy development

From Nearby Stars to Solar Kinematics: New Insight from Gaia DR2 Catalogue



The aim of this study is to analyse the FGK Main-Sequence nearby stars from Gaia DR2 catalogue to investigate the Solar kinematics.

Machine Learning (ML)

Machine learning is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to imitate how humans learn, gradually improving its accuracy.

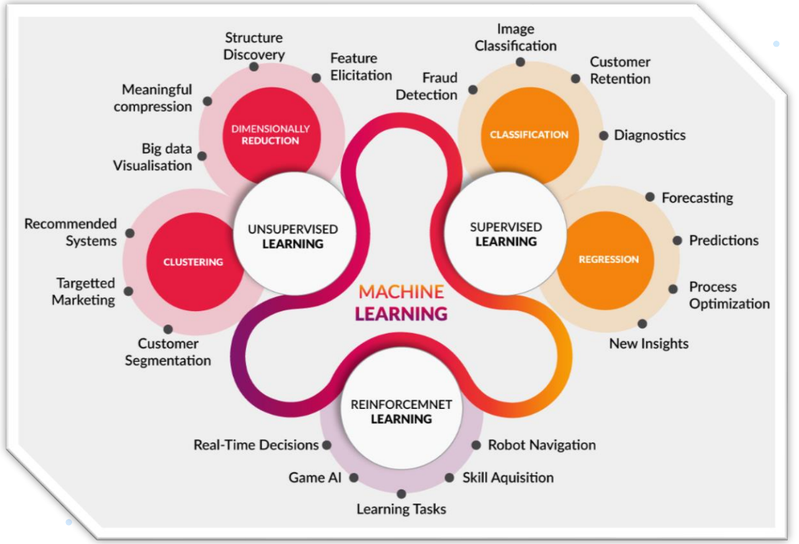
Supervised Learning

Supervised machine learning algorithms are designed to learn by example.

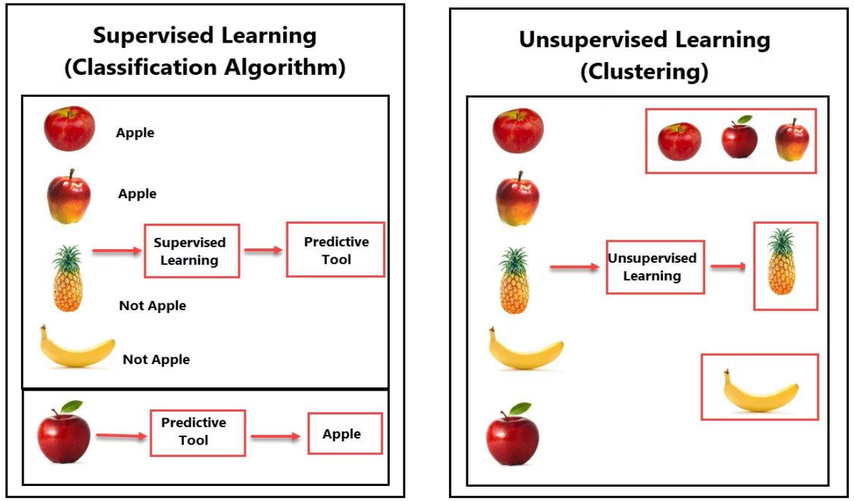
Unsupervised Learning

The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.

Machine Learning Diagram



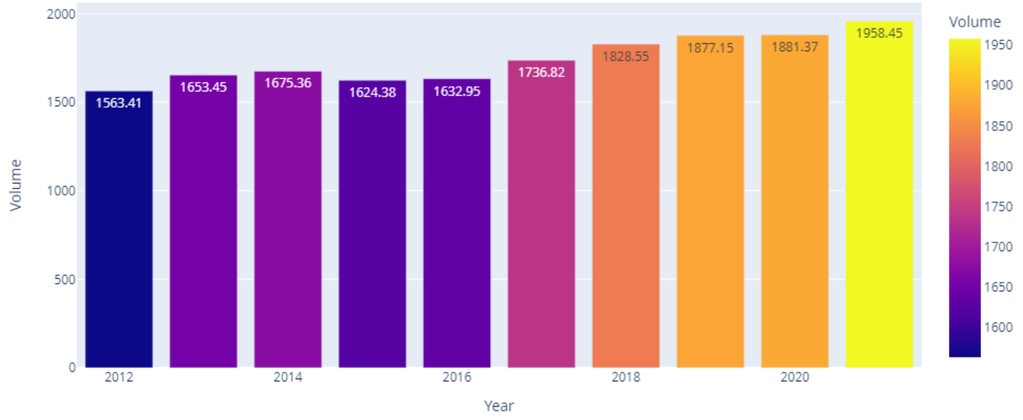
Supervised VS. Unsupervised



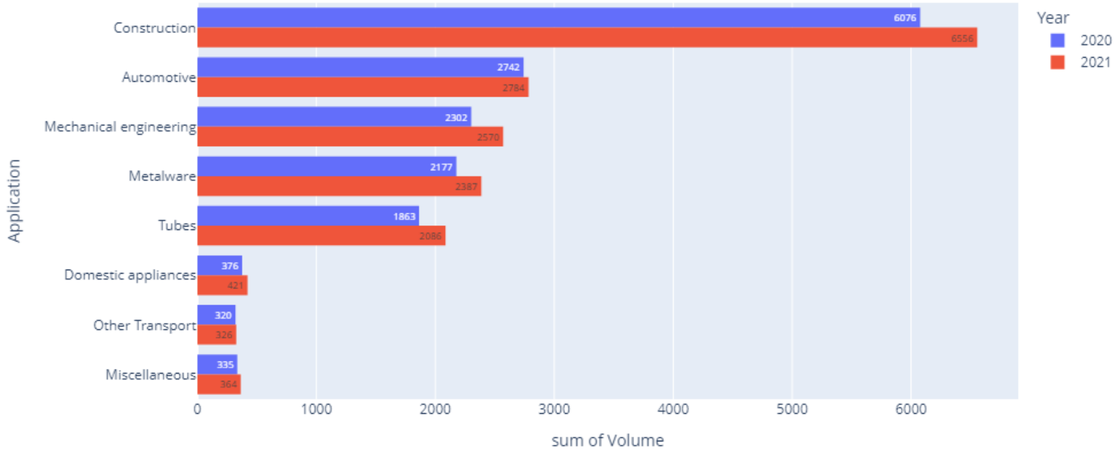
What is going on in steel industry?

Major steel-producing countries 2020 and 2021 (in million metric tons)

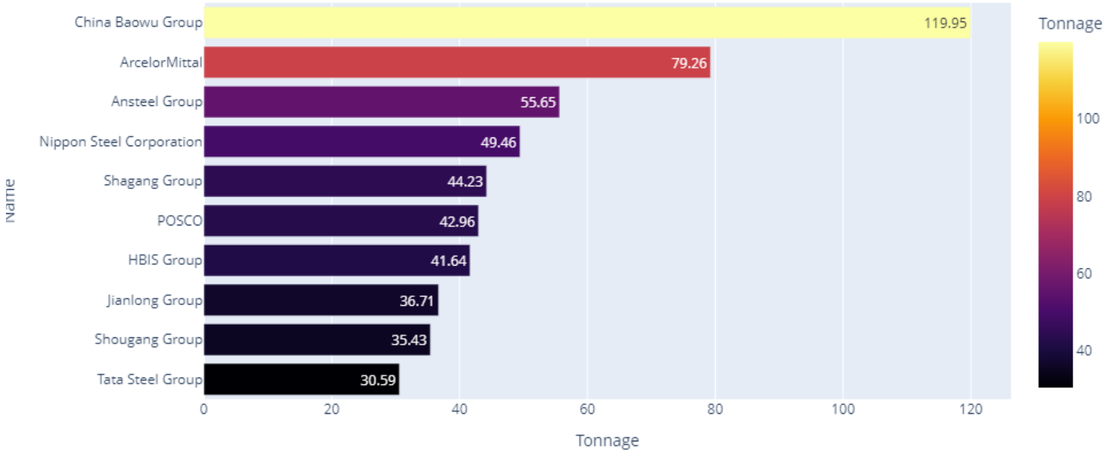
World crude steel production from 2012 to 2021 (in million metric tons)



EU Consumption: by sector of economic activity



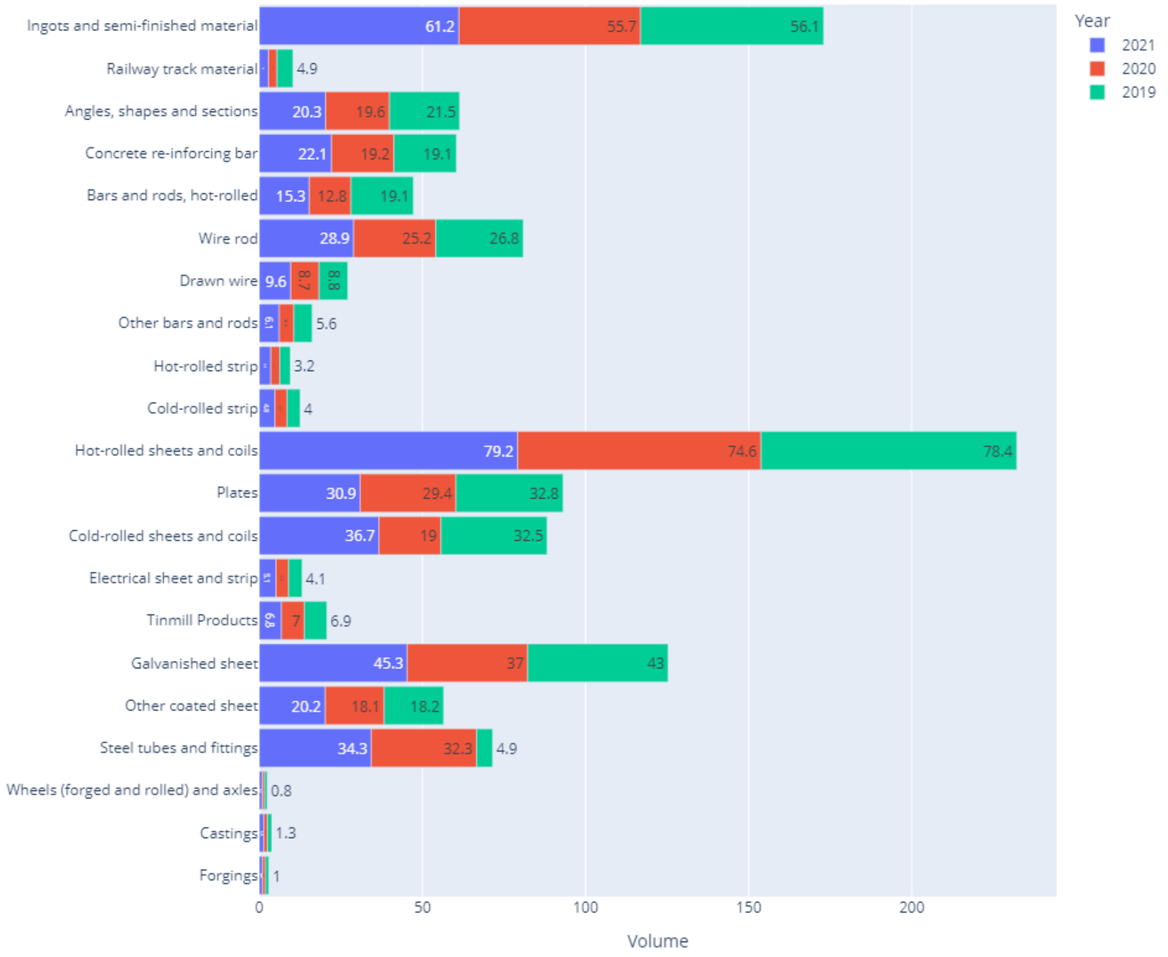
Top World steel-producing companies (in million metric tons)



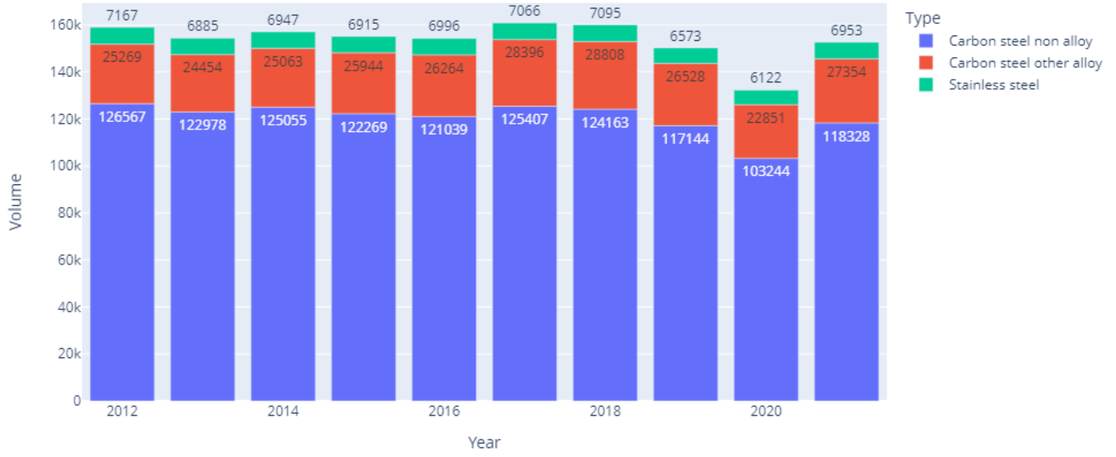
Data Sources: World Steel Association, Statista and UK Government

What is going on in steel industry?

World steel exports by product 2019 to 2021 (in million metric tons)



EU crude steel production: by quality



Description

The following are the most commonly used types of steel grades for hot rolling:

- A-36 (ASTM)
- 1010 (AISI)
- 1018 (AISI)
- A-1011 (ASTM)
- 1026 (AISI)
- A-500 (ASTM)
- 1045 (AISI)
- 1141 (AISI)

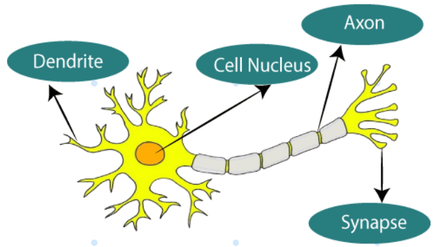
Data Sources: World Steel Association, Statista and UK Government

What are the correlations between chemical composition and properties?

Supervised Learning

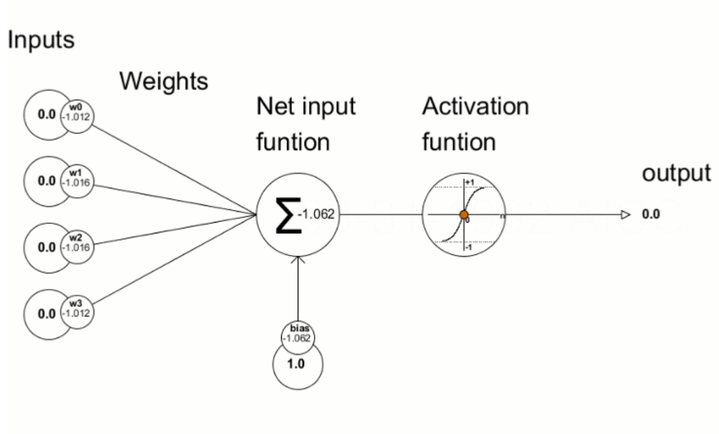
What is Artificial Neural Network?

The term "Artificial Neural Network" is derived from Biological neural networks that develop the structure of a human brain. Similar to the human brain that has neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks.



How does ANN work?

- The architecture of ANN consists of three layers.
- For the learning process, the data feds into the network via the input layer, and then the model learns by different combinations of input units, and consequently, it gives output.
- The neural network works by a feedback process called backpropagation. It means the expected outcome compares with the generated output and dedicates the weight by computing the differences between them. This process performs iteratively until the margin of error is minimal.



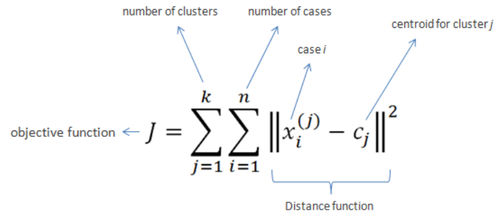
$$\sum_{i=1}^{n} Wi * Xi + b$$

How to find similarity and dissimilarity in different steel grades?

Unsupervised Learning

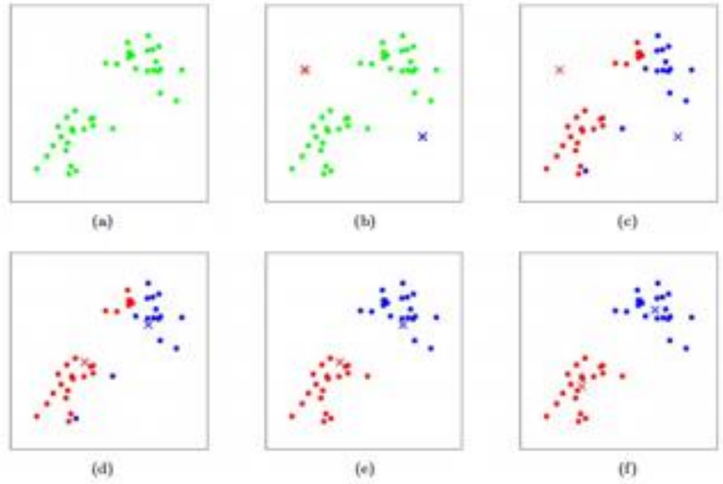
What is Clustering?

It automatically divides the data into clusters or groups of similar items. The clustering method aims to find meaningful patterns and structures in a data set.



What is K-Means?

1. Choosing the number of clusters: we must first determine the K number of clusters.
2. Initialising centroids: we choose random data points and identify them as the centroids for each group.
3. Assign each data point to the closest group: This step starts by calculating the distance between the centroid and data points.
4. Re-initialise centroids: The centroids will then be reset by averaging all the data points within that cluster.
5. Repeat steps 3 and 4: Steps 3 and 4 will be repeated until the optimal distance between the data points and their corresponding centroid is achieved.



Which parameter is important in each group and determines the group characteristic?

Explainable AI

SHAP or SHAPley Additive exPlanations

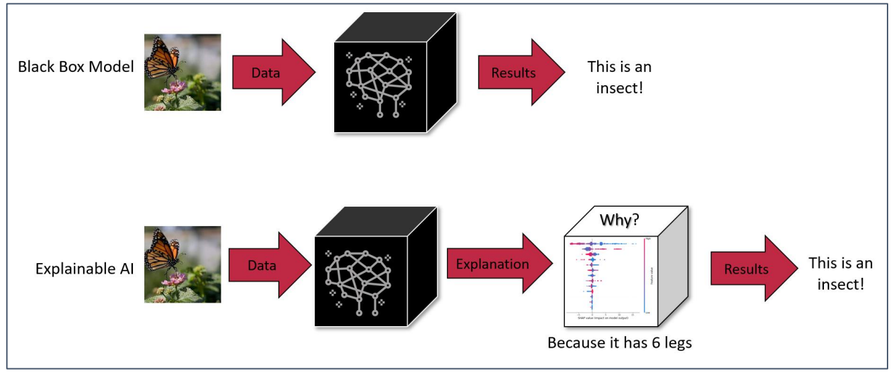
It is a visualisation tool that can be used for making a machine learning model more explainable by visualizing its output. It can be used for explaining the prediction of any model by computing the contribution of each feature to the prediction.

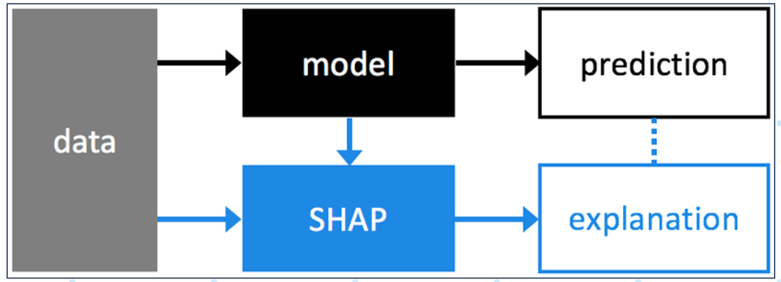
What is SHAP?

SHAP values are based on Shapley values, a concept coming from game theory. But game theory needs at least two things: a game and some players.

- the “game” is reproducing the outcome of the model,
- the “players” are the features included in the model.

What Shapley does is quantifying the contribution that each player brings to the game.





What have we got so far?

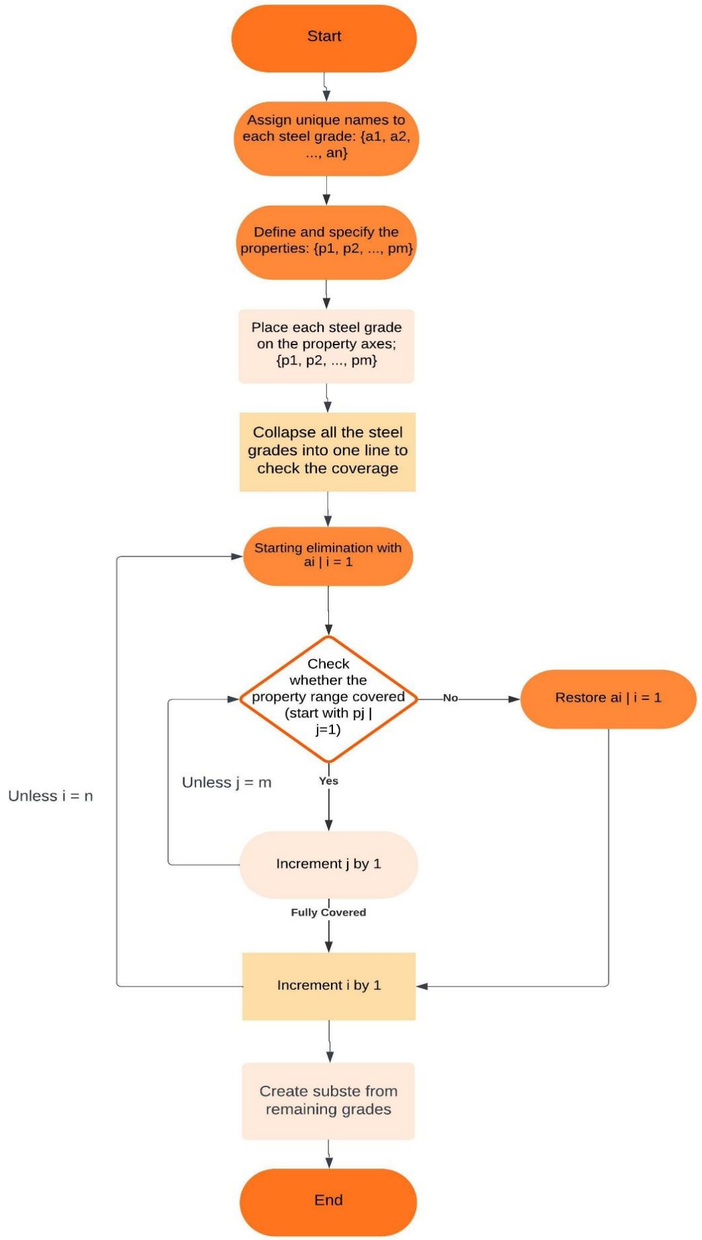
- Developed a comprehensive overview of the steel industry.
- We can identify relationships between chemical compositions and properties using ANN.
- We can determine the optimal number of clusters using the Elbow method and Silhouette score.
- We can identify similarities and dissimilarities using K-Means.
- We can detect overlaps within each generated group using dynamic bin histograms.
- We can uncover the characteristics of each generated group using SHAP.

We need an approach for reducing the number of steel grades based on the generated similar group.

Elimination Process

We propose an algorithmic elimination process to come up with the smallest number of grades within each generated group.

- The goal of this process is to determine which steel grades can be eliminated while ensuring the remaining steel grades cover the property ranges of those that are eliminated and have a minimum of recyclability-limiting elements (Tramp elements & Critical raw materials).



Steps of the project

Data Analysis

- Data Cleaning
- Data Scaling
- Dimension Reduction

1st Step

ANN

Supervised learning:
Finding relationship between chemical compositions and properties.

2nd Step

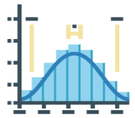
Clustering

Unsupervised learning:
Classify the data into groups based on similarities using K-Means.

3rd Step

Finding overlaps

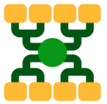
Create a histogram based on the range of properties to find characteristics of each generated group.



4th Step

Elimination

Algorithmic method to come up with the smallest number of grades within each generated group.



5th Step

Material

About the data

Data contained 270 rows and 43 attributes.

Attribute	Example	Data Type
Chemical Composition	C, Mn, P, S, …	Float/Int
Mechanical Properties	Hardness, UTS, YTS, Elongation	Float/Int
Treatment conditions	Quenching, Annealing, …	True/False
Standards	AISI, ASTM, EN	String

MatWeb

Physical Properties	Metric
Density	7.80 - 7.87 g/cc

Mechanical Properties	Metric
Hardness, Brinell	30 - 500
Hardness, Knoop	103 - 392
Hardness, Rockwell B	49 - 100
Hardness, Rockwell C	9.0 - 71
Hardness, Vickers	88 - 384
Tensile Strength, Ultimate	295 - 2210 MPa
Tensile Strength, Yield	165 - 1260 MPa
Elongation at Break	5.0 - 48 %
Reduction of Area	13.4 - 72.5 %
Modulus of Elasticity	186 - 210 GPa
Bulk Modulus	148 - 163 GPa
Poissons Ratio	0.29 - 0.30
Machinability	40 - 160 %
Shear Modulus	72.0 - 80.0 GPa
Izod Impact	3.00 - 135 J

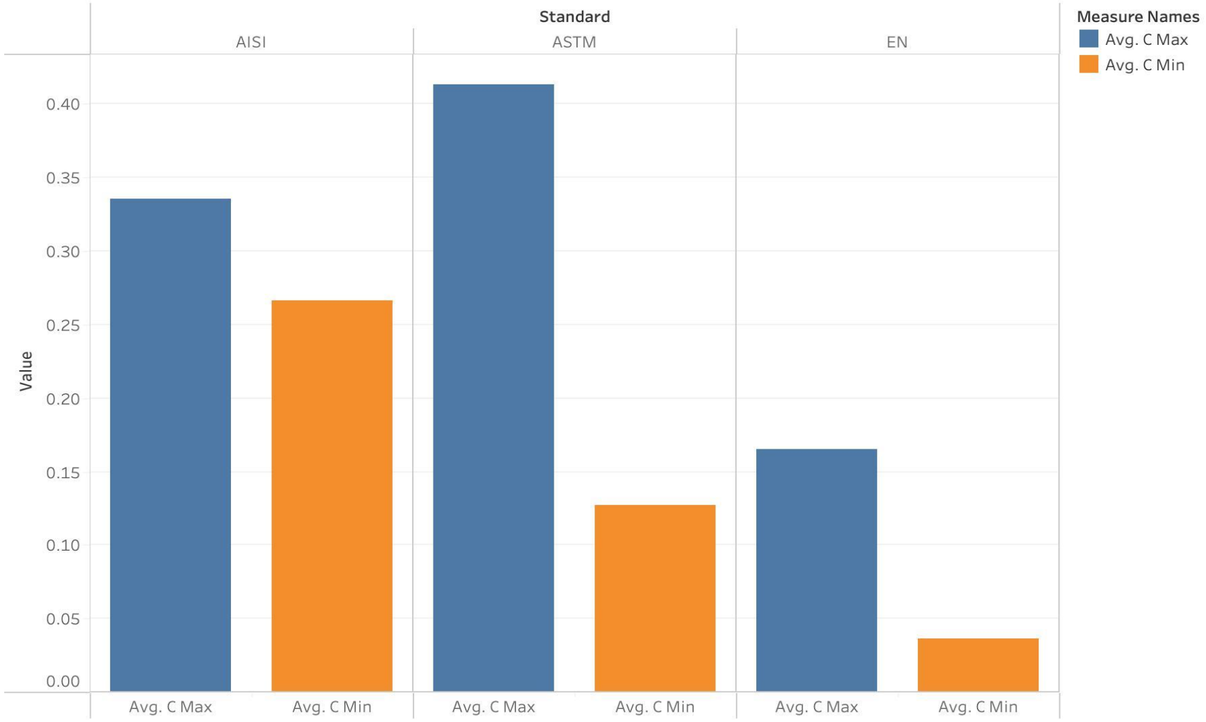
Electrical Properties	Metric
Electrical Resistivity	0.00000200 - 0.0000180 ohm-cm

Thermal Properties	Metric
CTE, linear	11.0 - 16.6 µm/m-°C
Specific Heat Capacity	0.470 - 0.490 J/g-°C
Thermal Conductivity	27.1 - 52.0 W/m-K
Melting Point	1430 °C
Transformation Temperature	270 - 760 °C

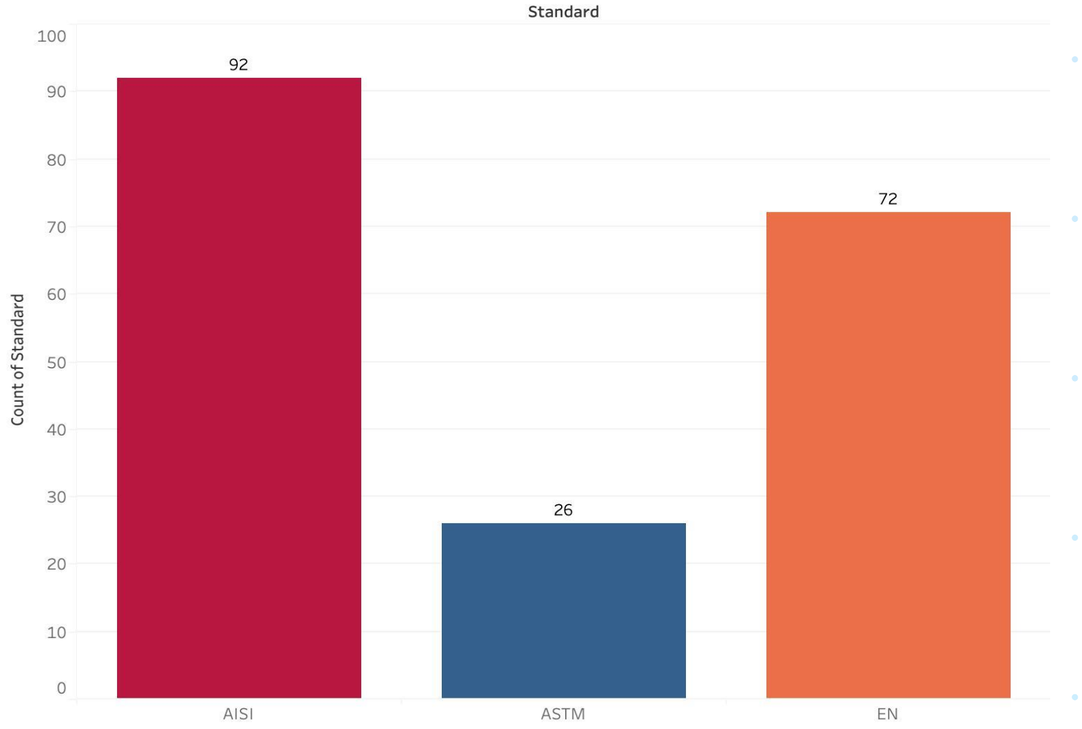
Component Elements Properties	Metric
Aluminum, Al	0.015 - 0.020 %
Carbon, C	0.060 - 1.03 %
Chromium, Cr	0.15 - 1.1 %
Copper, Cu	0.25 - 0.30 %
Iron, Fe	96.8 - 100 %
Lead, Pb	0.15 - 0.36 %
Manganese, Mn	0.25 - 2.05 %
Molybdenum, Mo	0.060 - 0.25 %
Nickel, Ni	0.10 - 0.40 %
Phosphorus, P	0.020 - 0.40 %
Silicon, Si	0.10 - 0.40 %
Sulfur, S	0.015 - 0.50 %
Titanium, Ti	0.0050 %
Vanadium, V	0.030 - 0.10 %

Data Analysis

Range of Carbon Content

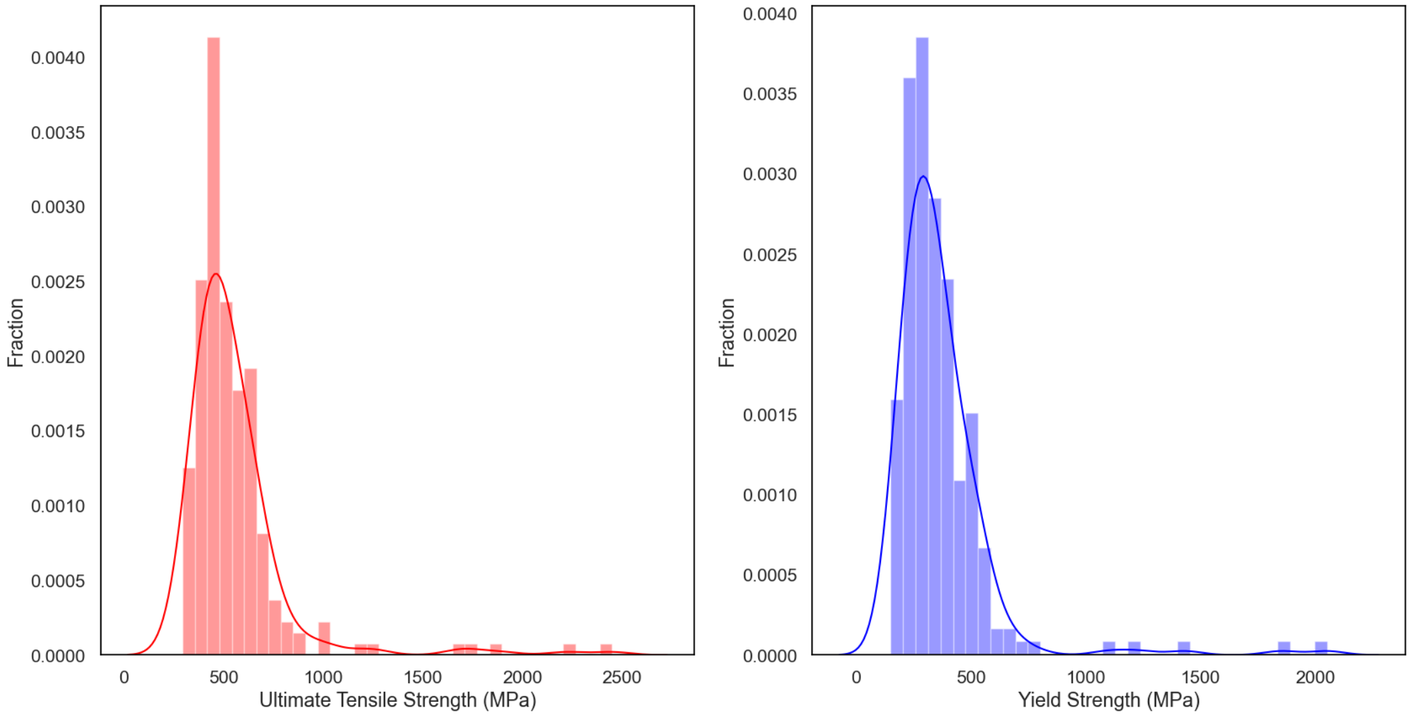


Distribution of Different Standards

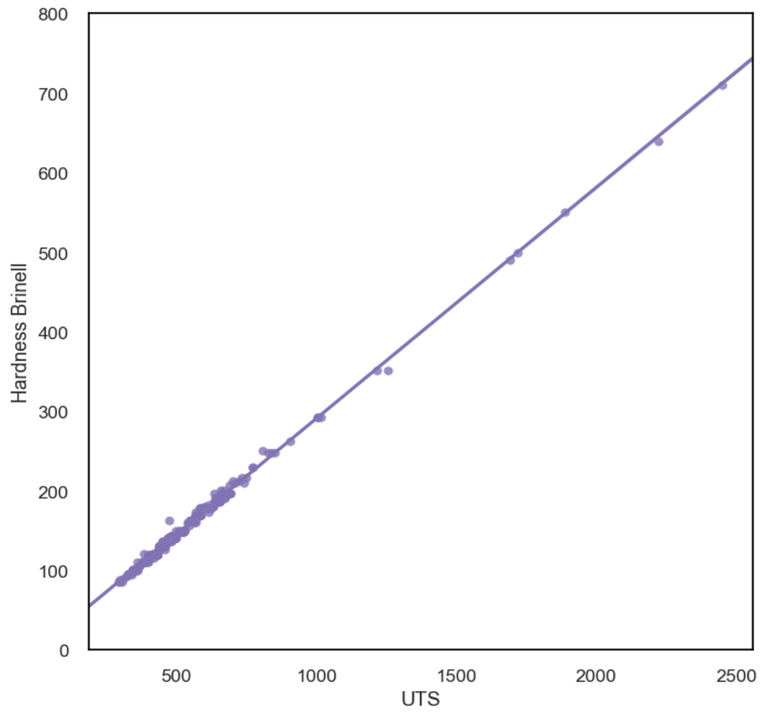


Data Analysis

Distribution of UTS and YS in the dataset

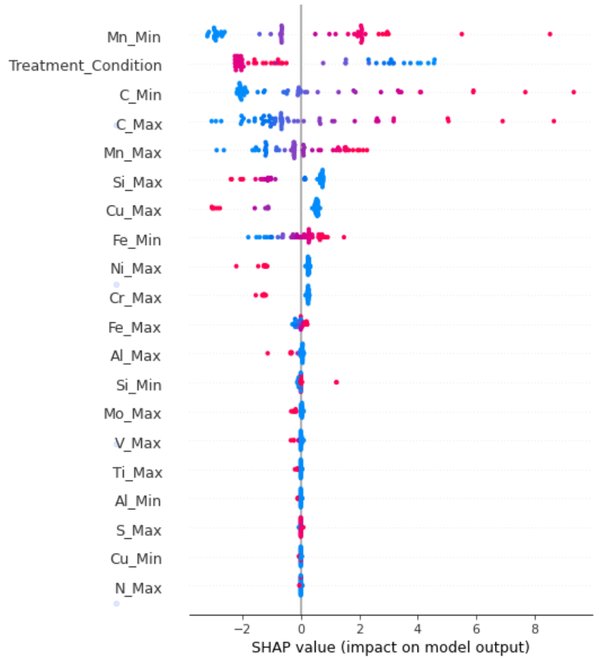


Correlation between UTS and Hardness

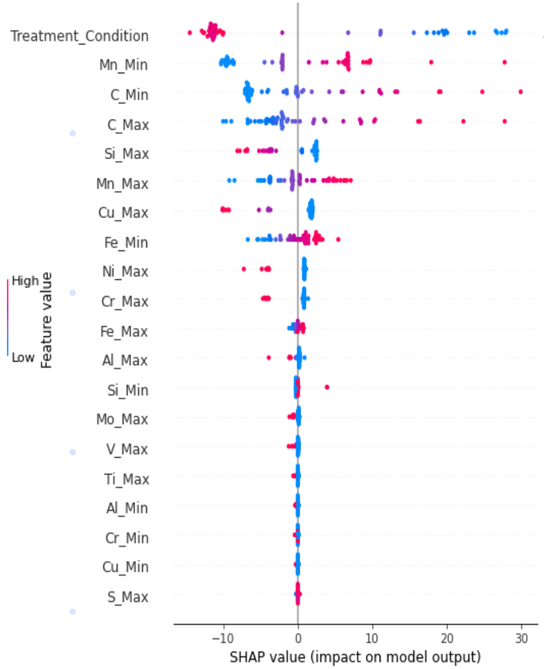


Relationship Between Chemical Elements and Properties

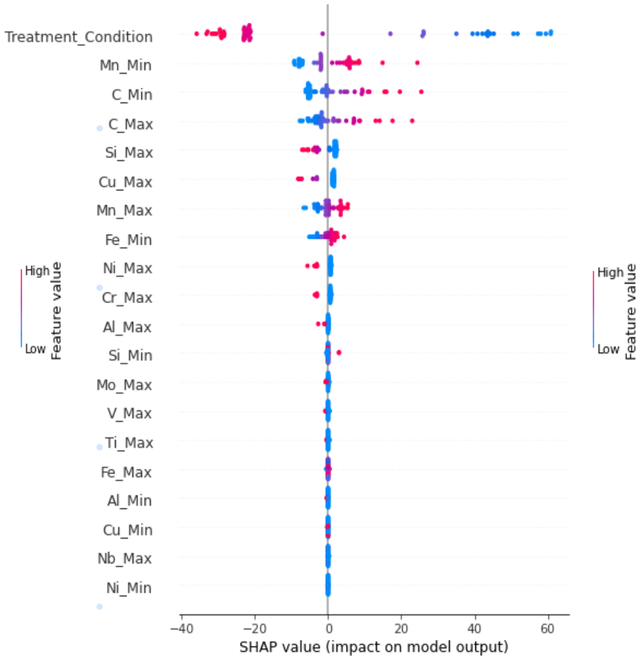
SHAP summary plot of hardness



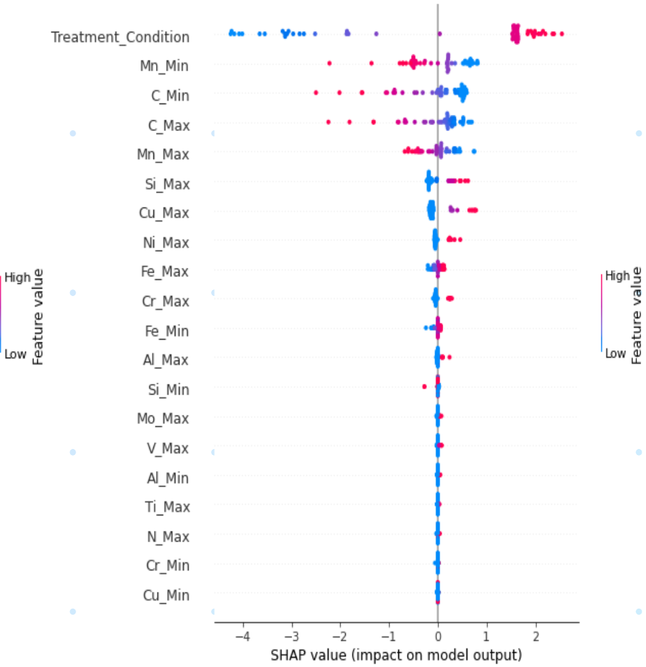
SHAP summary plot of UTS



SHAP summary plot of YS



SHAP summary plot of EL



Hardness Relationship

- Higher content of C and Mn positively influence hardness,
- Si and Cu content have a negative correlation.

UTS Relationship

- Higher content of Mn and C positively influence hardness,
- Si and Cu content have a negative correlation.

YS Relationship

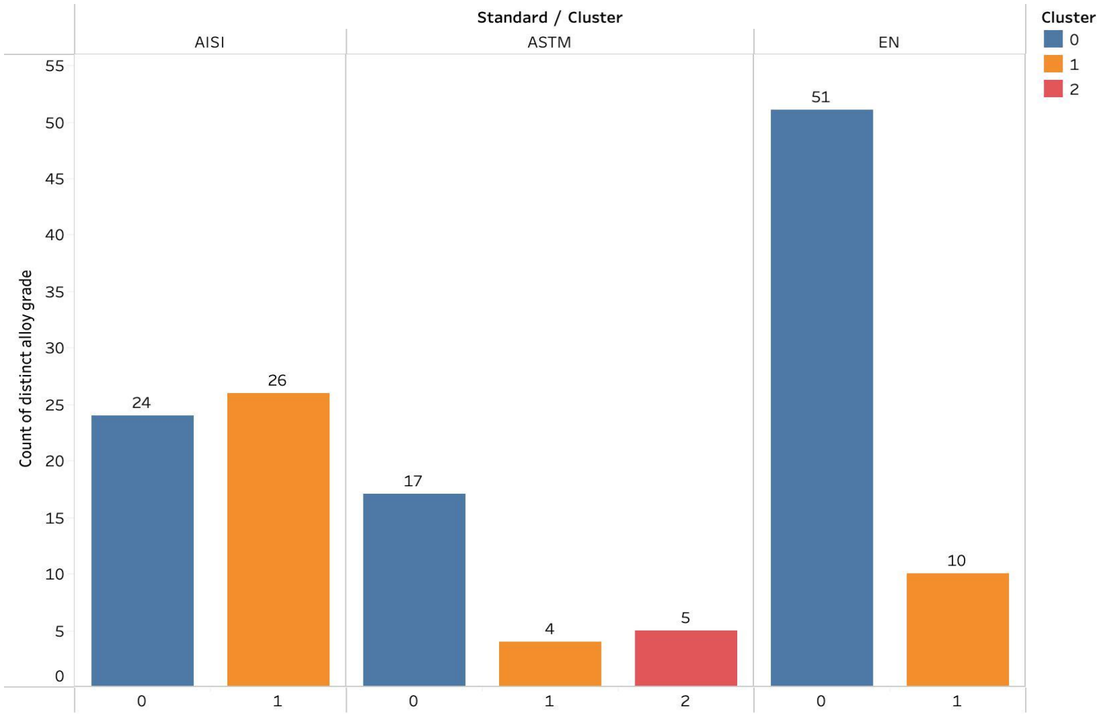
- Higher content of Mn and C positively influence hardness,
- Si and Cu content have a negative correlation.

Elongation Relationship

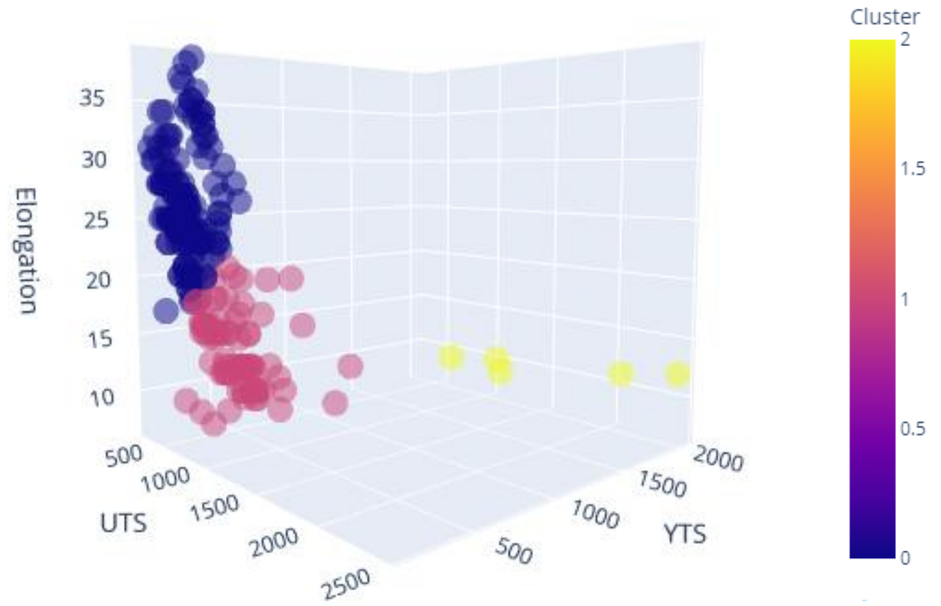
- Mn and C content have a negative correlation
- Higher content of Si, Cu, and Ni is associated with higher elongation

Finding Similarities and Dissimilarities

Generated group on Carbon steel data (Unsupervised Learning)



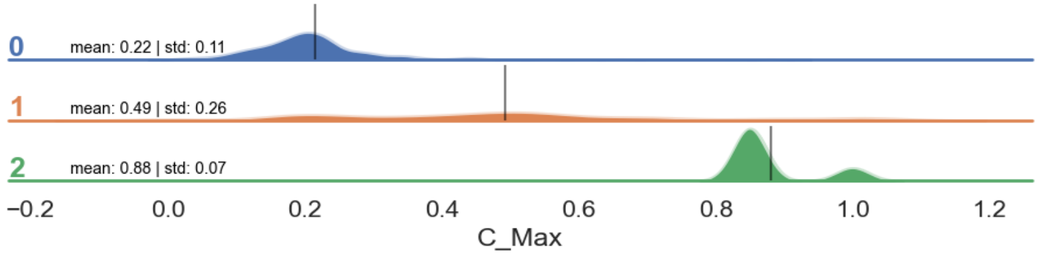
Clustering results using K-Means



Description

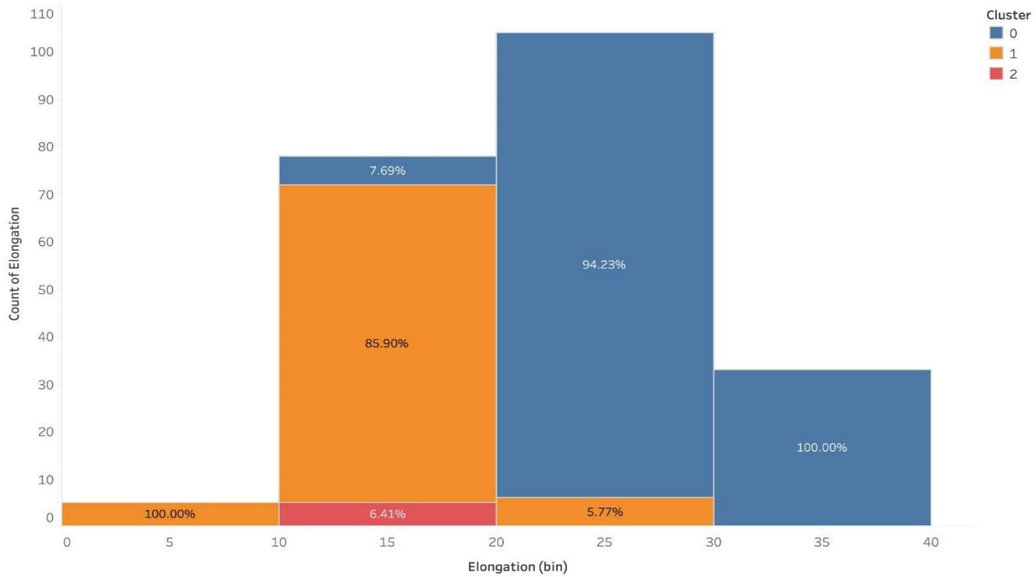
- 137 distinct steel gardes are classified into 3 different groups based on their mechanical properties.
- 92 grades of steel falling into group 0, 40 grades falling into group 1 and 5 grades of steel falling into group 2.

Distribution of carbon content

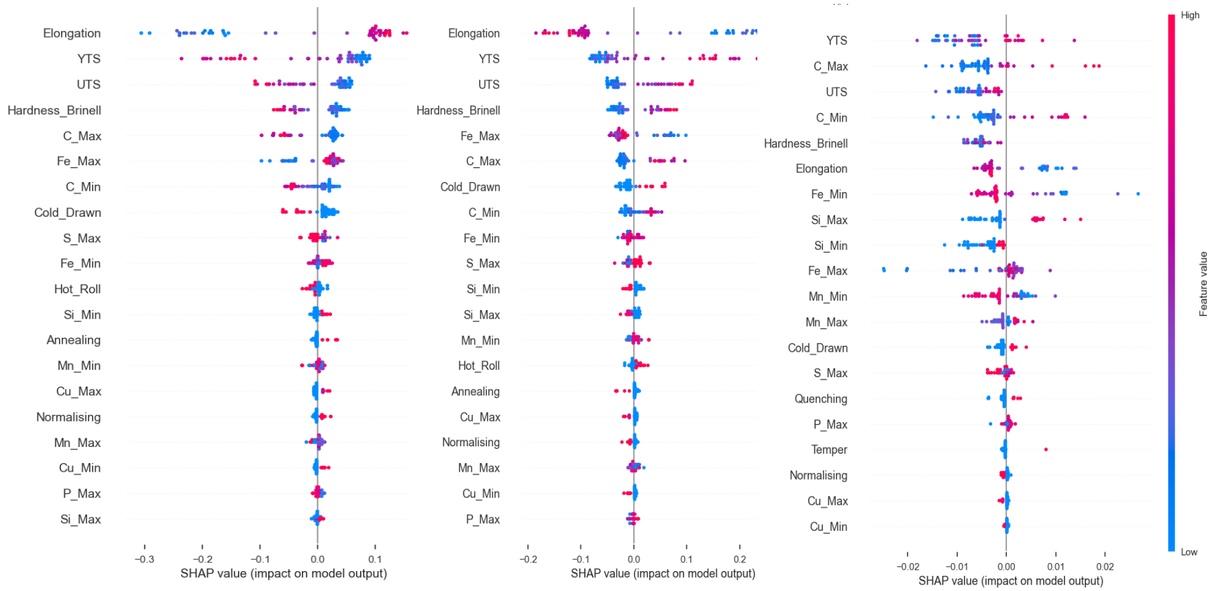


Finding characteristics of each group

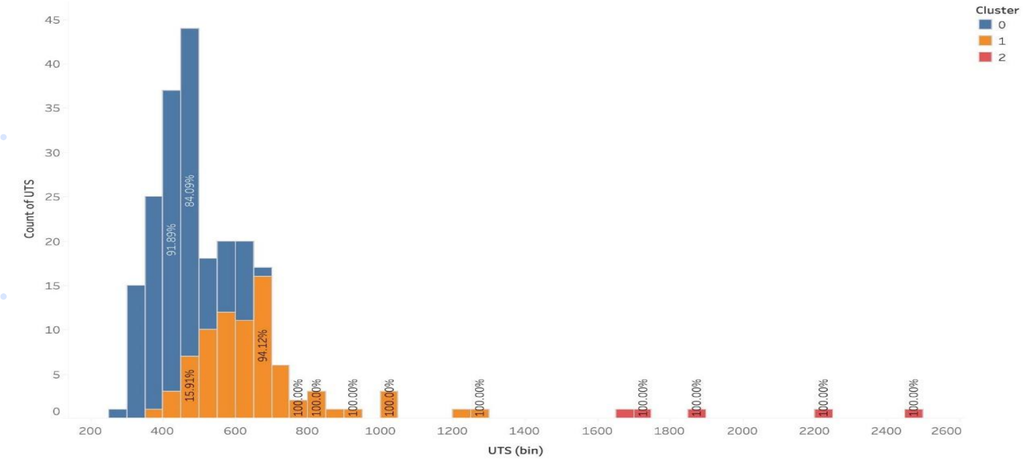
Elongation histogram (10% interval)



Clustering results using K-Means



UTS histogram (50MPa intervals)

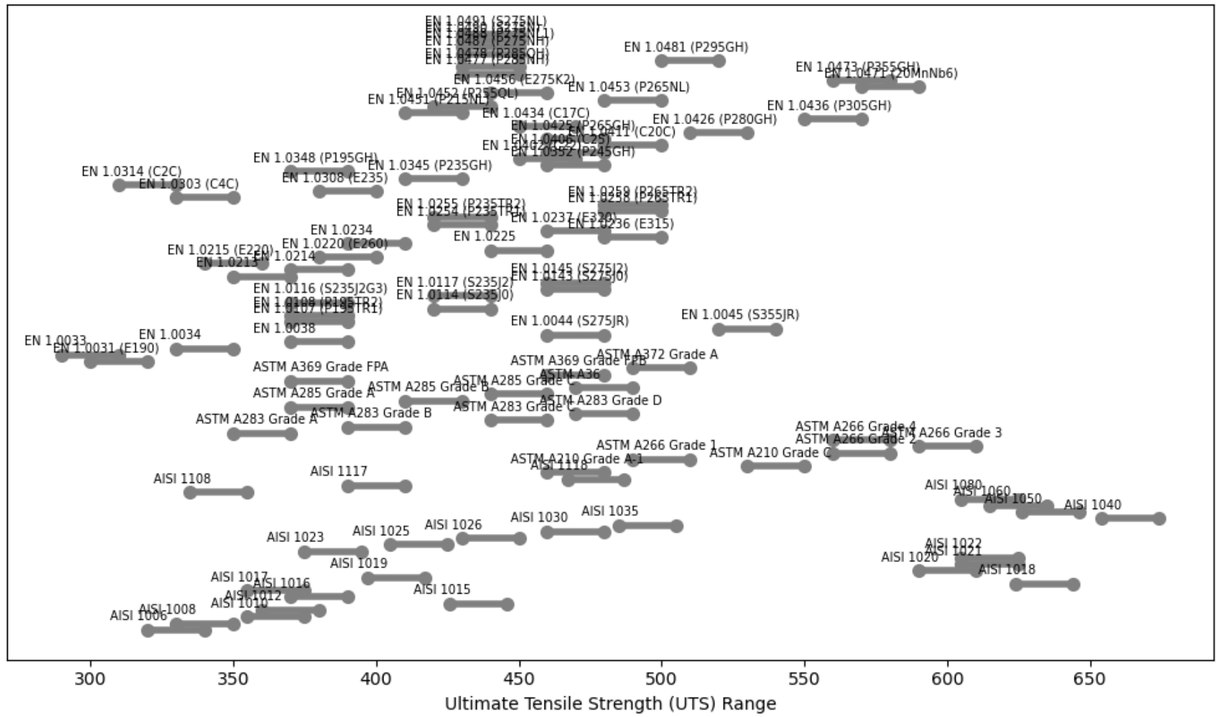


Description

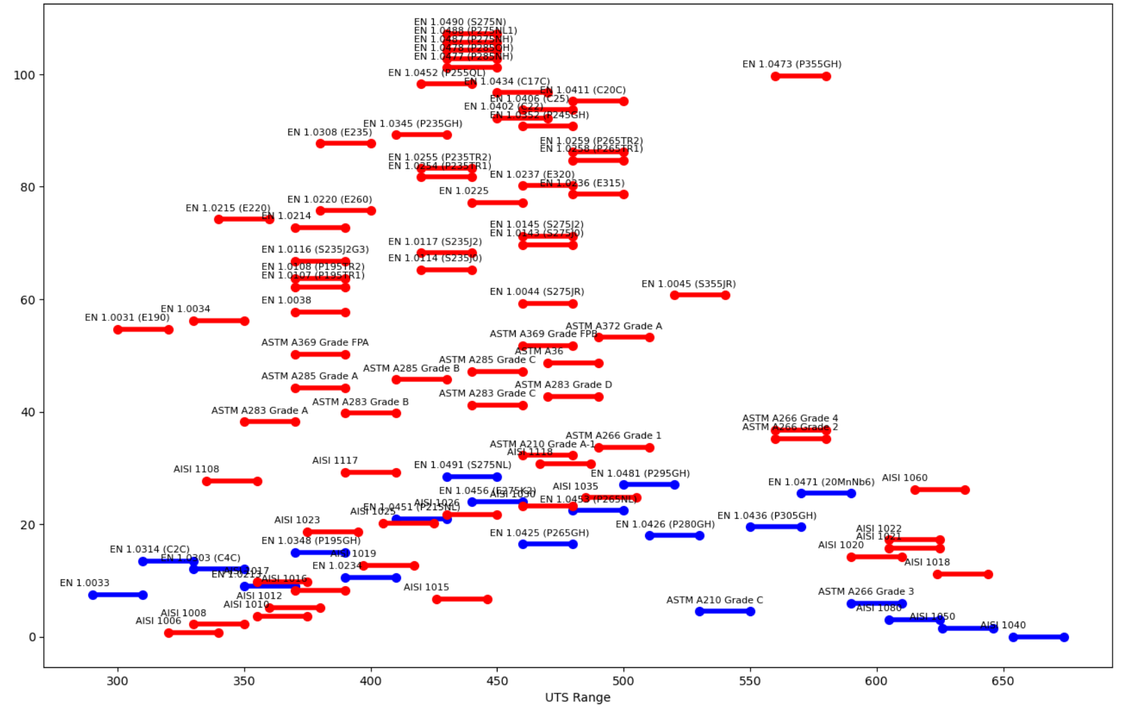
- Group 0: Well-suited for applications that require materials with high ductility and deformation.
- Group 1: Consists of alloy grades with moderate to lower elongation but with good strength properties.
- Group 2: Primarily comprises alloy grades with high strength and carbon content. well-suited for applications that demand superior strength and hardness properties.

Elimination Process – Group 0

Group 0 – Before Elimination on UTS axis



Group 0 – After Elimination on UTS axis



Remaining Grades

The remaining grades in group zero are: ['AISI 1040', 'AISI 1050', 'AISI 1080', 'ASTM A210 Grade C', 'ASTM A266 Grade 3', 'EN 1.0033', 'EN 1.0213', 'EN 1.0234', 'EN 1.0303 (C4C)', 'EN 1.0314 (C2C)', 'EN 1.0348 (P195GH)', 'EN 1.0425 (P265GH)', 'EN 1.0426 (P280GH)', 'EN 1.0436 (P305GH)', 'EN 1.0451 (P215NL)', 'EN 1.0453 (P265NL)', 'EN 1.0456 (E275K2)', 'EN 1.0471 (20MnNb6)', 'EN 1.0481 (P295GH)', 'EN 1.0491 (S275NL)']

Summary

- The ANN successfully identified correlations between properties and chemical composition, while SHAP was employed to offer insights on selected properties.
- The clustering method generated three distinct groups corresponding to the well-established groups of low, medium and high carbon steels based on mechanical properties only.
- The SHAP analysis successfully identified the most important features contributing to clustering and unrevealed the group characteristics.
- The elimination algorithm was tested and proved to be beneficial in reducing the number of existing steel grades.
- This algorithm can be used in different areas of science and technology for the purpose of reducing the options or eliminating any undesirable entities from a classification.

Future Work

- To ensure the reliability and applicability of the proposed reclassification methodology, it is crucial to expand the dataset to include a broader range of steel grades, including stainless steel and other alloys.
- Including details about how steel grades are used in real-world applications will add significant value to the elimination process.

SYMPOSIUM ON FERROUS METALLURGY 2024

Thank You!

Any Question?

Rationalisation of steel grades and specifications

Email: Sadegh.Jalalian@brunel.ac.uk

This work was supported by UK Research and innovation (UKRI)
under the CircularMetal Centre program.

Brunel University London | UKRI Interdisciplinary Centre for Circular Metals | UK Research and Innovation | NICER PROGRAMME | BCAST