

Sustainability Classification for Fashion Products



Janey



TABLE OF CONTENTS

01 BACKGROUND

02 METHODOLOGY

03 EDA

04 MODEL
DEVELOPMENT

05 INDEPENDENT
DATASET

06 FUTURE WORK



RISE OF THE ECO-CONSCIOUS CONSUMER

2019

35%

Consumers chose sustainable products to help protect the environment

2021

74%

37%

Consumers looked for products with environmentally friendly packaging

78%

42%

Consumers would pay more for sustainably produced items

80%

PROBLEM STATEMENT

FASHION INDUSTRY

Produced **2-8%** of global carbon emissions
Cause of microplastics in oceans

GREENWASHING

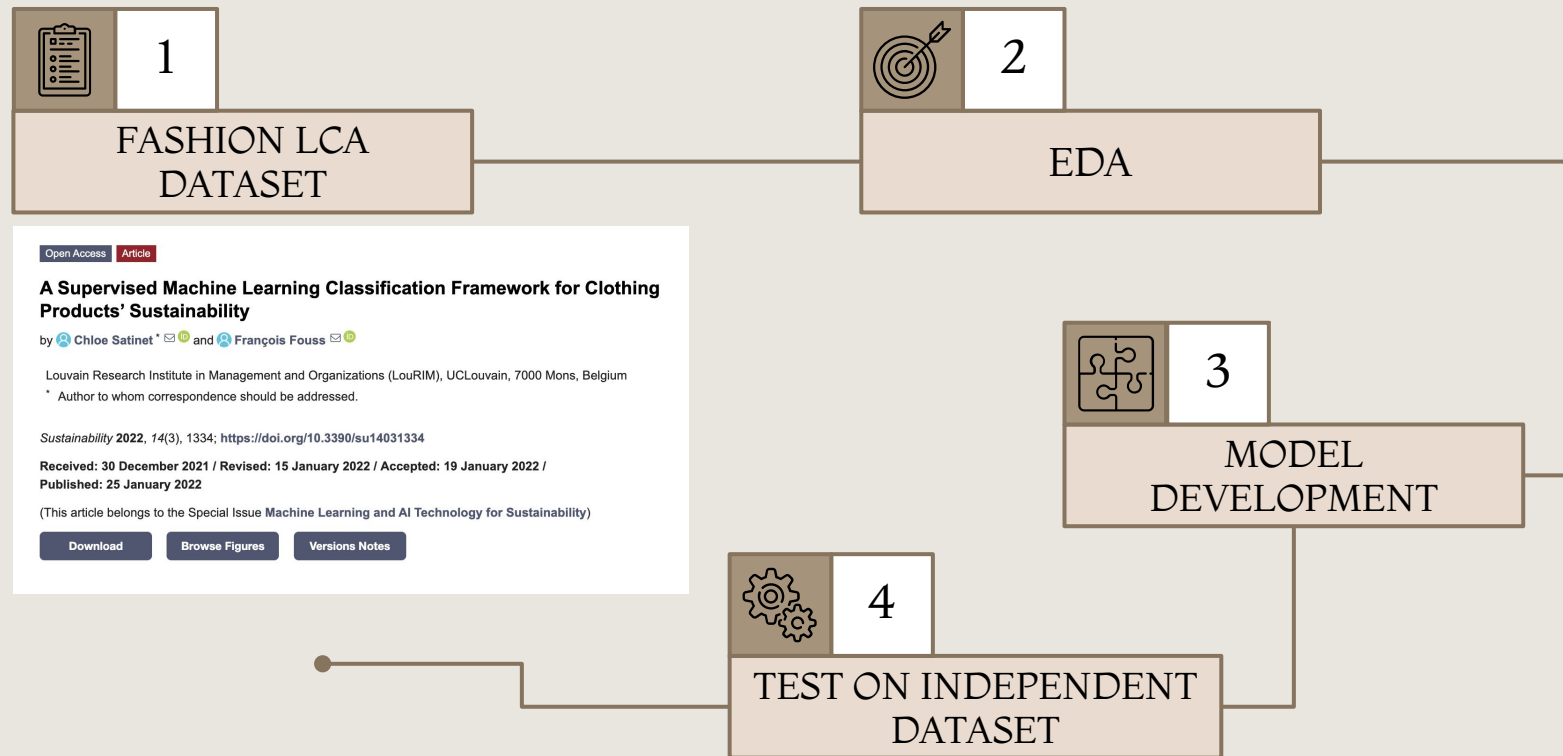
Misleading, vague
claims that their
products are
sustainable

CERTIFICATIONS

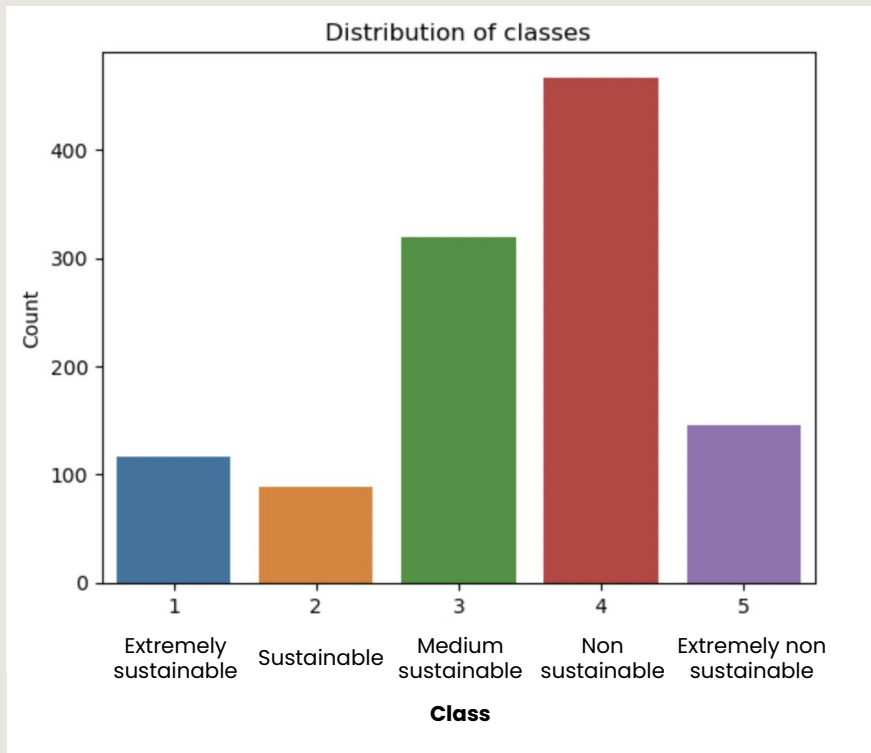
Expensive and
inaccessible to small
businesses



METHODOLOGY



EDA – Distribution of target variable



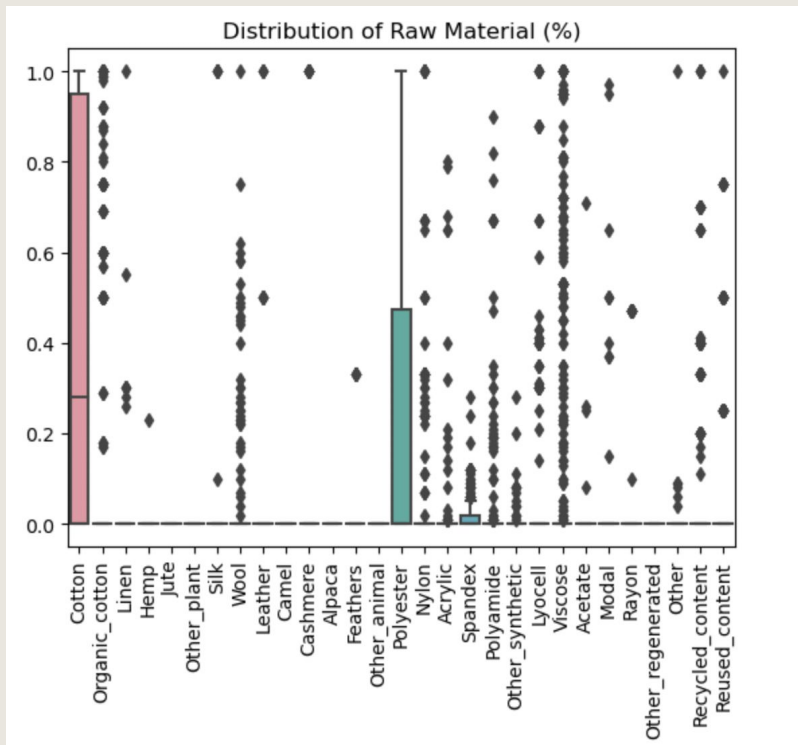
- Imbalanced classes
 - F1-score
- Multi-class
 - Macro average

EDA – Significance of categorical features

FEATURES	CHI-SQUARE P-VALUE
Material label	2.667e-97
Chemical label	5.589e-16
Production label	3.709e-125
Reusability label	2.464e-135
Recyclability label	1.687e-126

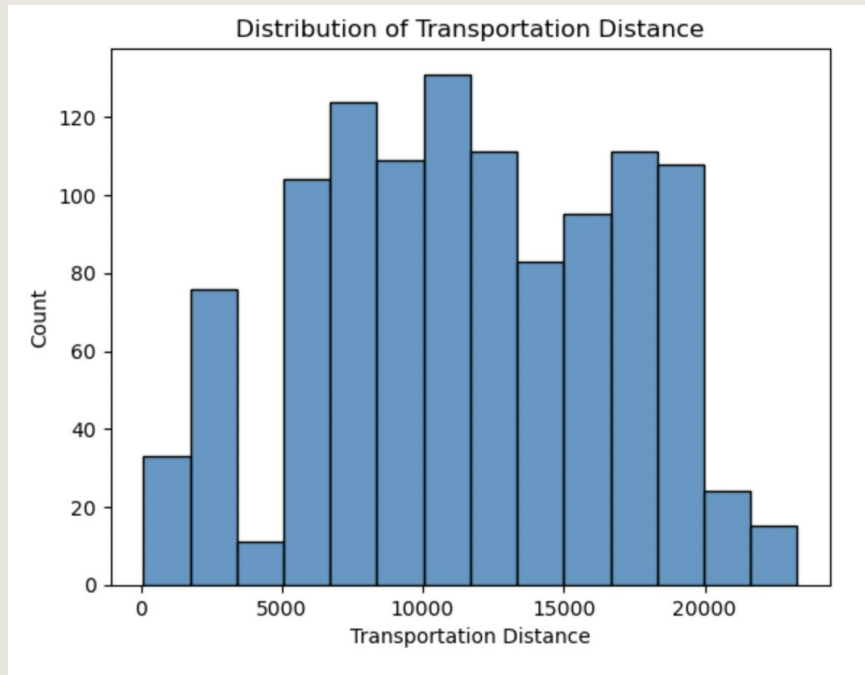
- Categorical vs categorical
 - Chi-square test
- p-values < 0.05
 - Statistically significant
- Keep all features

EDA – Distribution of raw material features



- Many outliers
- ML model that can deal with outliers

EDA – Distribution of transportation distance



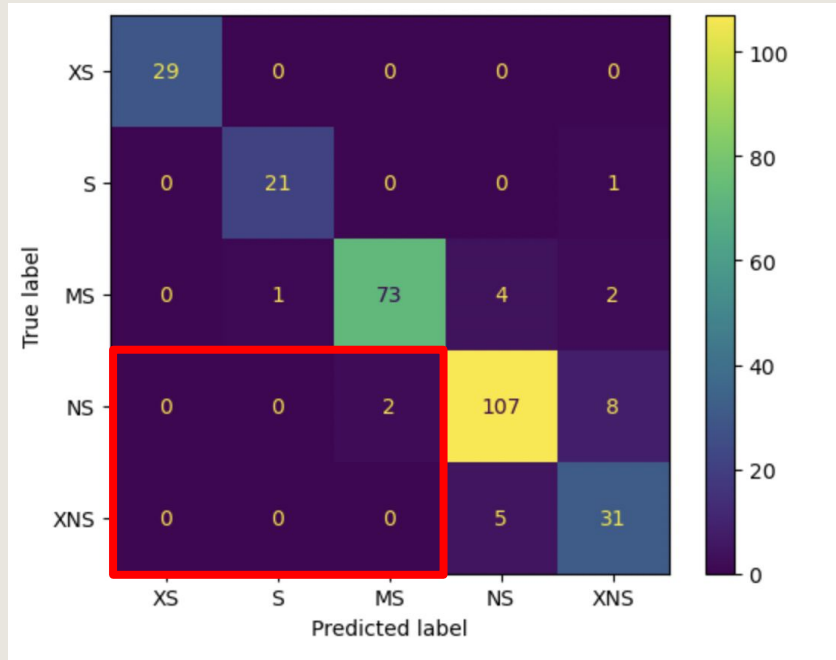
- One outlier
- Non-normal distribution
- Non-parametric ML model



MODEL DEVELOPMENT

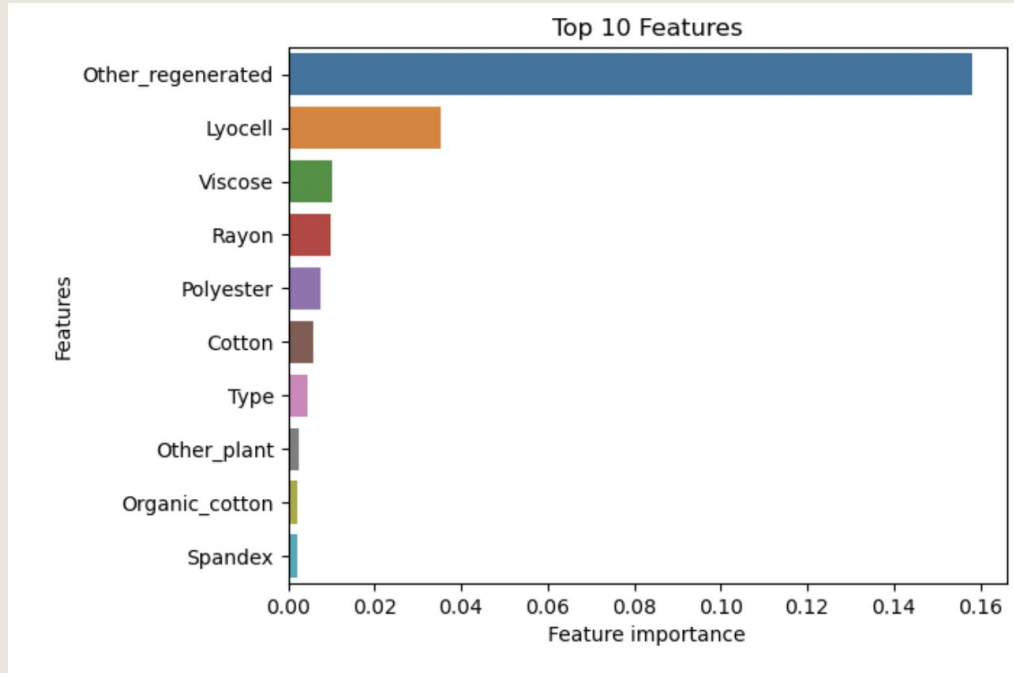
MODELS	TRAIN F1	TEST F1	REMARKS
K-Nearest Neighbors	0.9986	0.8661	Overfitted
Support Vector Machine	0.8891	0.8893	Convergence issues
Random Forest	0.942	0.922	Good fit
Decision Trees	0.9436	0.9172	Good fit
Gradient Boost	0.9916	0.9267	Overfitted

CONFUSION MATRIX



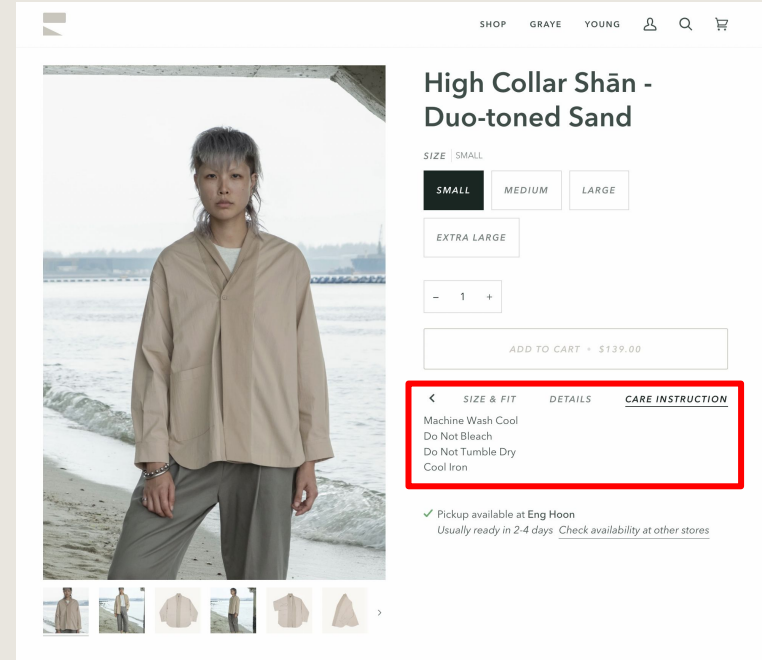
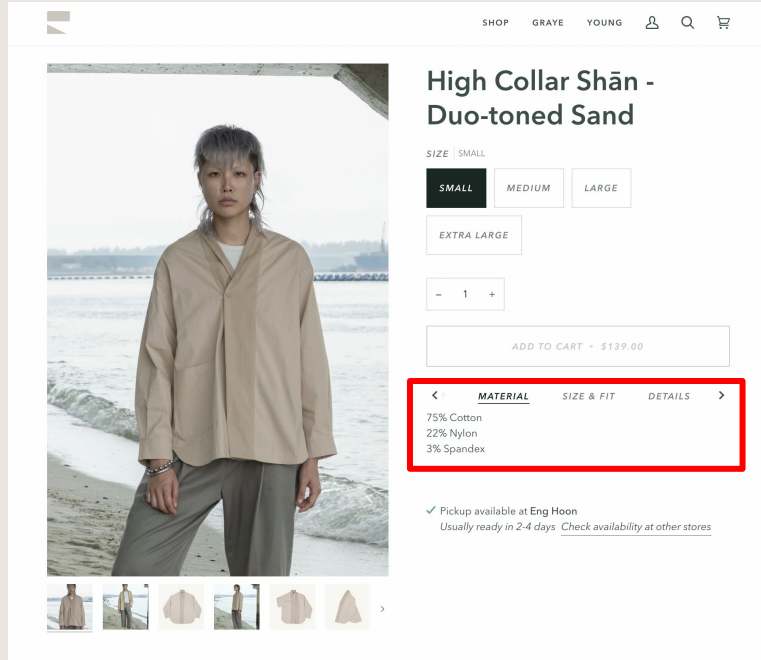
- Minimise errors in the red box
- Good model

FEATURE IMPORTANCE



- Raw materials contribute the most to the classification

INDEPENDENT DATASET – Data collection



TEST ON INDEPENDENT DATASET

		Cotton	Organic_cotton	Other_plant	Wool	Other_animal	Polyester	Nylon	Spandex	Polyamide
		mean	mean	mean	mean	mean	mean	mean	mean	mean
Type	EI									
jacket	1	0.000000	0.0	1.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0
	3	0.937500	0.0	0.000000	0.0	0.0	0.000000	0.057500	0.003750	0.0
shirt	1	0.000000	0.0	1.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0
	3	0.748000	0.0	0.000000	0.0	0.0	0.000000	0.048000	0.014667	0.0
short	3	0.830000	0.0	0.000000	0.0	0.0	0.075000	0.095000	0.000000	0.0
sweater	3	1.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0
t-shirt	3	0.730000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.020000	0.0
trousers	3	0.657826	0.0	0.086957	0.0	0.0	0.026087	0.152174	0.003913	0.0

- Did not just predict majority class

- Natural > synthetic
- Linen > Cotton

Model good enough to be deployed



FUTURE WORKS

MODEL IMPROVEMENT

- Random oversampling
- Feature engineering
- Other independent testing

SUSTAINABILITY CONTEXT

- Dashboard for companies to check on their sustainability initiatives
- Extend the model for other product categories

THANKS

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution

