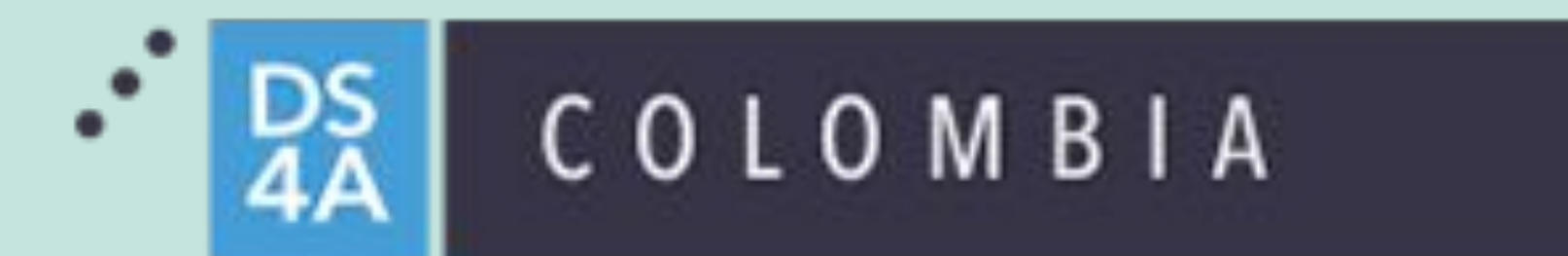


PREDICTING ACADEMIC DROPOUT AT UNIVERSIDAD DE CÓRDOBA



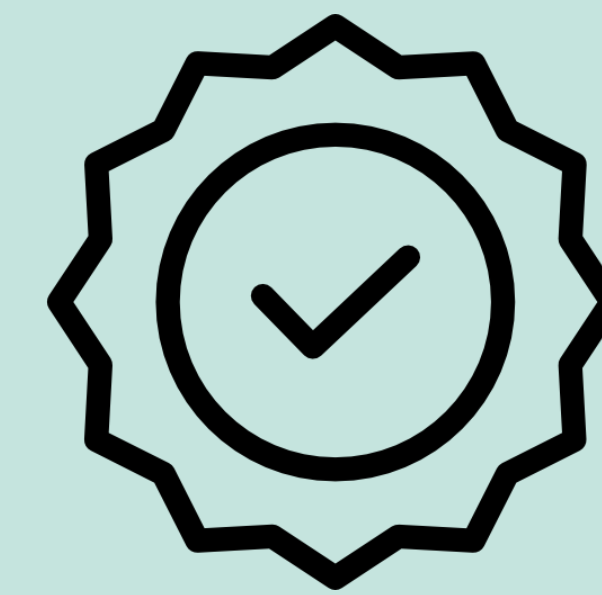
Team 64 - Fabio Bombiela, Dennis Garcia, Ivan Molina, Daniel Otero, Juan Rodríguez, Erika Rojas, Carolina Albarracin

Highlights

- Interactive dashboard to monitor Moodle's platform usability.
- Despite of the pandemic, Universidad de Cordoba does not have many academic dropout problems compared to the academic dropout figures in Colombia.

Background

By receiving government funding due to its public nature, Universidad de Cordoba seeks to benefit the Colombian population by providing quality education. Therefore, by anticipating student desertion, it avoids the loss of economic resources and endorses its institutional reputation.



Data



PowerCAMPUS

Information provided by Universidad de Cordoba contains .csv files from the Moodle platform, a learning management system, and PowerCAMPUS platform, the official academic information system of Unicordoba. The dataset shared contains::

- Student and Teacher usage logs of Moodle platform, that includes login activity, course navigation, activities, assignments and resources' usage.
- Socioeconomic and academic information of students, including dropout per course and semester.

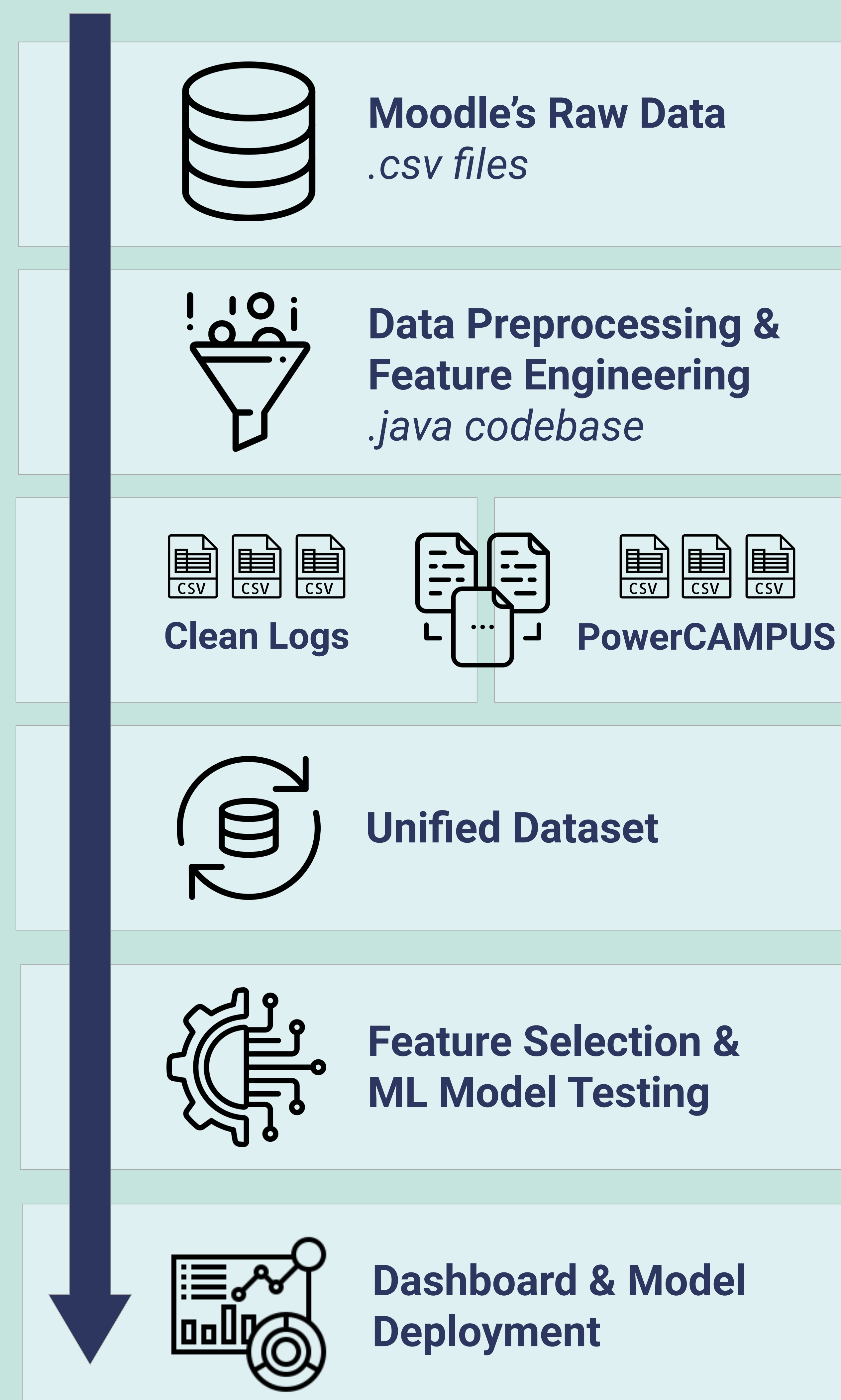
+93M Moodle's usage logs

+16K Unique students

~18GB Raw Data



Data Processing



Model

Features such as gender, socioeconomic level, number of courses and logs at moodle's platform, directly impacts dropout.



The following models were applied to classify students more prone to dropout: logistic regression, decision tree, random forest classifier and XGBoost classifier.

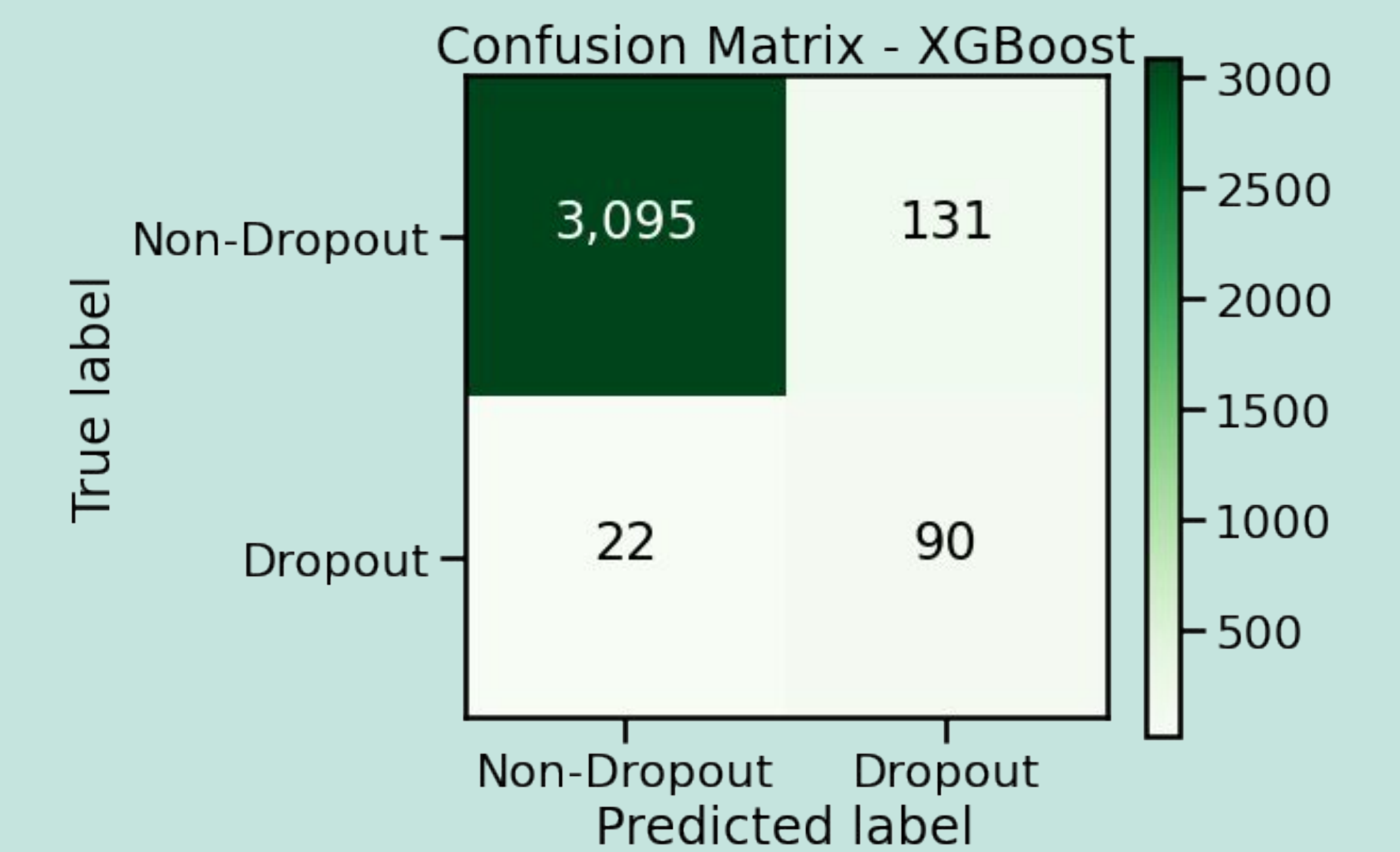
For model assessment, metrics such as precision, recall, F1 and accuracy were taken into account to determine overall model performance.

Model	Metric Accuracy	Recall	F1	Precision
Logistic Regression	84,51%	83,04%	25,51%	13,09%
Decision Trees	77,11%	89,29%	20,75%	11,74%
Random Forest Classifier	86,28%	83,93%	29,10%	11,99%
XGBoost Classifier	95,42%	80,36%	54,05%	40,72%

Finding out that **XGBoost Classifier** delivers better prediction results overall, based on the assessed metrics.

Results & Deployment

Despite of heavy class imbalance, the machine learning model is capable of classifying students prone to dropout effectively.



The classification model was deployed along with an interactive dashboard, in which you can visualize Moodle platform usability statistics at different levels of granularity. Additionally, it has the ability to classify students, at course level, with higher chances of dropout.

STUDENT'S GENDER	NUMBER OF LOGS	NUMBER OF COURSES
FEMALE	15	5
CURRENT SEMESTER	SOCIOECONOMIC LEVEL	1ST CUTOFF GRADE AV.
SECOND SEMESTER	2	2.98

CALCULATE RISK

STUDENT AT DROPOUT RISK!

Database & Backend

Frontend

