**Group**

**Request 1: Topic and research question**

A close up of text

AI-generated content may be incorrect.

A close-up of a paper

AI-generated content may be incorrect.

I think we have an example here (See the highlight).

**Request 2: Dataset**



A text on a page

AI-generated content may be incorrect.

I believe we clearly mentioned the biases (yellow part) and challenges (green part) in my report.

**Request 3: Discussion**

For this section, I think we compare those 3 model comprehensively, we even create a table to clearly illustrate it. Would you mind to check it again?

**Individual – lshe0103**

**Request 1: Predictive model - Model description**

A close-up of a text

AI-generated content may be incorrect.

A text on a page

AI-generated content may be incorrect.

I have already provided an example. Specifically, I noted that this model is capable of identifying hidden relationships between the attributes, which directly addresses our research question. In addition, I pointed out that it supports one-hot encoding, which is consistent with the categorical variables in our dataset. These examples illustrate how the model is well-suited to our data characteristics and research goals.

**Request 2: Predictive model - Model algorithm**

A grey rectangular sign with black text

AI-generated content may be incorrect.

A close-up of a text

AI-generated content may be incorrect.

In the yellow section, I believe I have described the hyperparameters concisely, including both what they are and how they affect training. In the green section, I have mentioned most of the commonly used activation functions in recent years for MLPs. Therefore, I do not think there are anything else I can provide.

**Request 3: Predictive model - Model development**

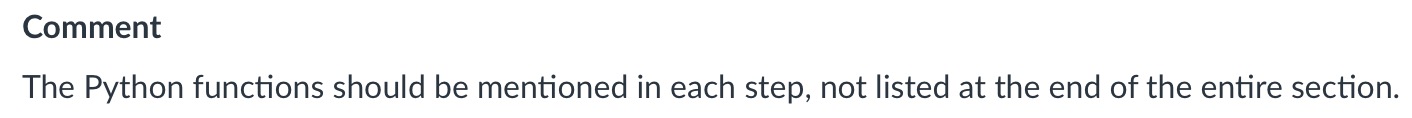


This comment seems rather vague. However, I believe I have described the building process thoroughly enough to meet the HD-level criteria. I have detailed the feature engineering methods I used, the hyperparameters I selected and the rationale behind those choices, as well as the implementation code I provided. Could you please clarify which specific aspects you feel are missing, based on the HD-level criteria?

**Individual – qzho0498**

**Request 1**

**Predictive model - Model development**



In this section, I aimed to integrate function mentions and implementation details directly within each explanation step to maintain a smooth narrative flow. For example, during feature preprocessing, I described how `StandardScaler` was applied via `ColumnTransformer`, and in the training phase, how `eval\_set` and `early\_stopping\_rounds` were used with `xgboost.XGBClassifier`. These functions were not listed at the end as a block, but rather woven into each corresponding paragraph where their functional role was explained. I believed this format provided more clarity than simply appending a code summary afterward.

The model development pipeline in this project mainly involved standard scikit-learn and XGBoost utilities. As a result, describing them inline, along with their roles (e.g., `GridSearchCV` for tuning, `PredefinedSplit` for fixed validation).Each preprocessing and training step was explicitly connected to its Python implementation. For instance:

- Standardization: `StandardScaler` inside `ColumnTransformer`

- One-hot preservation: default `passthrough` behavior

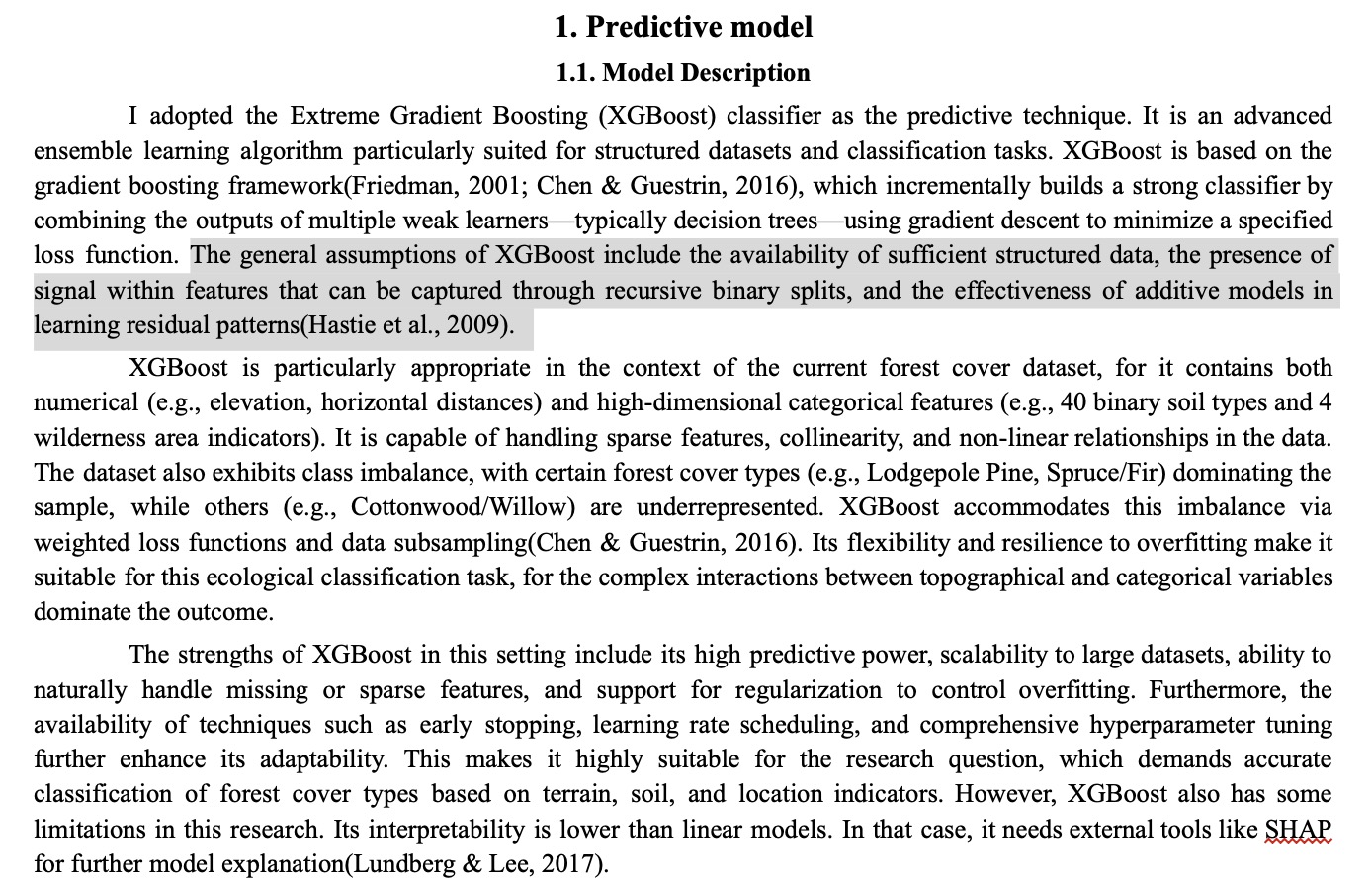
- Training monitoring: `eval\_set`, `early\_stopping\_rounds=50`

- Hyperparameter search: `GridSearchCV` and `RandomizedSearchCV`

Other functions used are not the core modules, so I did not list them in the limited report.

**Request 2**

图片包含 文本

AI 生成的内容可能不正确。

Since the assumptions for the entire study have been described in detail in group component 1, in model assumption, I list the common assumptions of the XGBoost model.