Data Science Ethics Checklist

A. Data Collection

A.1 Informed consent:
If there are human subjects, have they given informed consent, where subjects affirmatively opt-in and have a clear understanding of the data uses to which they consent?
The raw data were derived from prior published studies involving wet lab experiments on pharmaceutical removal from wastewater using biochar. As no human subjects were involved, obtaining informed consent was unnecessary.

A.2 Collection bias:
Have we considered sources of bias that could be introduced during data collection and survey design and taken steps to mitigate those?
Bias was minimized during data collection by including literature from four prominent journal databases: Google Scholar, Scopus, PubMed, and Web of Science.

A.3 Limit PII exposure:
Have we considered ways to minimize exposure of personally identifiable information (PII) for example through anonymization or not collecting information that isn't relevant for analysis?
No personally identifiable information (PII) was gathered as part of this project, so the risk of exposing such data does not exist.

A.4 Downstream bias mitigation:
Have we considered ways to enable testing downstream results for biased outcomes (e.g., collecting data on protected group status like race or gender)?
As no data on human subjects were collected, addressing downstream bias was not applicable to this project.

B. Data Storage

B.1 Data security:
Do we have a plan to protect and secure data (e.g., encryption at rest and in transit, access controls on internal users and third parties, access logs, and up-to-date software)?
The data are stored on team members' personal drives and in a GitHub repository for wider accessibility. There are no plans to restrict or secure access to this data.

B.2 Right to be forgotten:
Do we have a mechanism through which an individual can request their personal information be removed?
This is not relevant to the project, as no personal information was collected.

B.3 Data retention plan:
Is there a schedule or plan to delete the data after it is no longer needed?
There is no plan to delete the data, as it may be required for future use.

C. Analysis

C.1 Missing perspectives:
Have we sought to address blindspots in the analysis through engagement with relevant stakeholders (e.g., checking assumptions and discussing implications with affected communities and subject matter experts)?
Input from team members and experts from the I-GUIDE initiative has been critical in identifying and addressing potential blind spots in the analysis.

C.2 Dataset bias:
Have we examined the data for possible sources of bias and taken steps to mitigate or address these biases (e.g., stereotype perpetuation, confirmation bias, imbalanced classes, or omitted confounding variables)?
The dataset was reviewed for bias during both the collection and analysis phases. Multiple databases were used during collection, and irrelevant variables were excluded during Exploratory Data Analysis (EDA) and preprocessing.

C.3 Honest representation:
Are our visualizations, summary statistics, and reports designed to honestly represent the underlying data?
All visualizations, summary statistics, and reports have been created to accurately and honestly reflect the data.

C.4 Privacy in analysis:
Have we ensured that data with PII are not used or displayed unless necessary for the analysis?
This is not relevant to the project since no PII data were collected or used.

C.5 Auditability:
Is the process of generating the analysis well documented and reproducible if we discover issues in the future?
The analysis process has been well-documented on GitHub and the I-GUIDE platform, ensuring reproducibility in the event of future concerns.

D. Modeling

D.1 Proxy discrimination:
Have we ensured that the model does not rely on variables or proxies for variables that are unfairly discriminatory?
Careful selection of variables ensures that the model does not depend on discriminatory variables or their proxies.

D.2 Fairness across groups:
Have we tested model results for fairness with respect to different affected groups (e.g., tested for disparate error rates)?
Model performance has been tested using a separate dataset to verify fairness across different groups.

D.3 Metric selection:
Have we considered the effects of optimizing for our defined metrics and considered additional metrics?
The potential impacts of optimizing for selected metrics, as well as the consideration of additional metrics, were addressed throughout the project.

D.4 Explainability:
Can we explain in understandable terms a decision the model made in cases where a justification is needed?
The Jupyter Notebook contains detailed notes explaining the model's decisions. Additionally, a Graphical User Interface (GUI) is being developed to provide clearer insights for a broader audience.

D.5 Communicate bias:
Have we communicated the shortcomings, limitations, and biases of the model to relevant stakeholders in ways that can be generally understood?
The project team has had detailed discussions about the model's limitations and biases. These will be communicated to stakeholders during the preparation of scientific manuscripts.

E. Deployment

E.1 Redress:
Have we discussed with our organization a plan for response if users are harmed by the results (e.g., how does the data science team evaluate these cases and update analysis and models to prevent future harm)?
The dataset assists decision-makers in identifying optimal conditions for pharmaceutical removal using biochar. While the model's predictions are not guaranteed to be 100% effective, users bear the responsibility for their decisions based on the outputs. The data science team cannot be held liable for any discrepancies.

E.2 Roll back:
Is there a way to turn off or roll back the model in production if necessary?
Yes, the model can be turned off or reverted if needed.

E.3 Concept drift:
Do we test and monitor for concept drift to ensure the model remains fair over time?
Regular updates to the dataset with newly published studies would be essential to maintain fairness and relevance. However, no immediate plans exist to update the dataset, which could impact the model's long-term applicability.

E.4 Unintended use:
Have we taken steps to identify and prevent unintended uses and abuse of the model and do we have a plan to monitor these once the model is deployed?
The primary objective is to make the knowledge widely accessible. The dataset and model will be available on GitHub to support user decision-making without plans for monitoring unintended uses.