BDM 3014 - Introduction to Artificial Intelligence

Proposal Submission - Lab 1 (Group 4)

AI in Predictive Toxicology

STUDENT'S NAME	STUDENT'S ID
Andres Santa	C0931978
Johan Rodriguez	C0931102
Amandeep Singh	C0937432
Amita	C0935607
Manjot Kaur	C0938094
Anju Bala	C0935847
Sarita Rani	C0938516
Brian Martinez	C0940439

The report should consist of ->

#1 How your industry/domain works (don't make it very generic, try to be speciDomain: The pharmaceutical industry, with a focus on safety and medication development.

Domain: The pharmaceutical industry, with a focus on safety and medication development.

Industry Function: Before new medications are put on the market, this sector is in charge of creating, testing, and verifying their efficacy and safety. In this procedure, predictive toxicology—which uses artificial intelligence (AI) to forecast drug toxicity—is essential because it aids in the early detection of any adverse responses.

Current Patterns: Predictive toxicology powered by AI is starting to be used to analyze chemical structures and forecast unfavorable medication responses. This change is driven by the goal to minimize late-stage failures in drug research in order to cut costs and speed up schedules.

Tools: Machine learning algorithms that analyze trends in chemical structures and databases like Tox21, ToxCast, and PubChem, which contain vast amounts of chemical and biological data, are used by researchers.

#2 Problem or opportunity the project is addressing (define the problem in a clear, specific & measurable way, who is the target consumer, why should the target consumer care about solving this problem) - 20 pts,

Problem: Toxicological problems found in late testing phases result in high failure rates in medication development, raising expenses and causing delays.

Solution: The industry may lessen the possibility of expensive late-stage failures by utilizing AI to anticipate possible toxicities early on, guaranteeing that only safer medications proceed through the development pipeline.

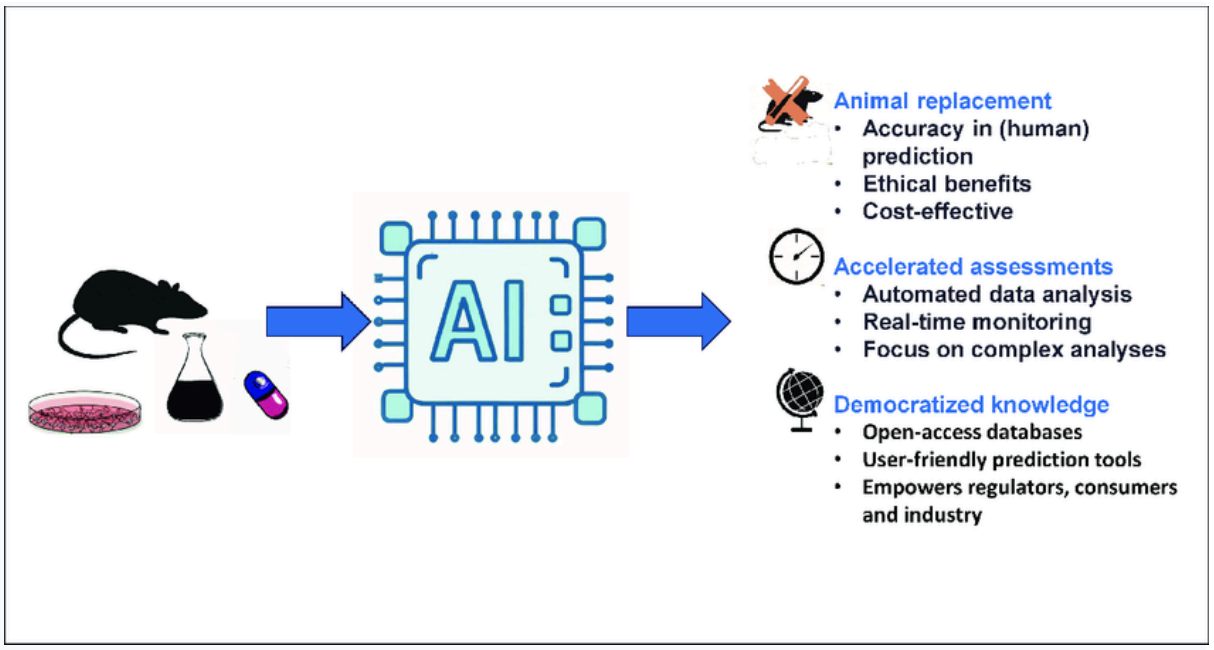
Pharmaceutical companies, drug developers, and healthcare organizations are the target consumers.

Impact on customers: By identifying toxicities early on, money can be saved, less animal testing can be done, and eventually, customers can get safer medications.

#3 Differentiation (what other solutions are available, why should they use your product as opposed to other solutions available elsewhere) - 15 pts

Uniqueness: AI-driven predictive toxicology forecasts toxicity based on chemical structures using data from several databases, in contrast to traditional toxicology, which frequently depends on animal testing or in-vitro experiments. This method can yield insights earlier in the drug development process and is quicker.

Benefit: Predictive toxicology using AI can lessen reliance on expensive and time-consuming lab testing. Additionally, it provides a higher level of accuracy in early-phase toxicity prediction, enabling businesses to prioritise the creation of safer substances.



#4 Data Gathering (Dataset(s) you are working on, how are you planning to gather the data, what does the data look like) - 20 pts and

For this project, three main datasets are used: Tox21, ToxCast, and PubChem. Each dataset provides valuable information related to chemical toxicity.

Tox21 includes data on chemical toxicity that is linked to various biological pathways.

ToxCast provides information on thousands of chemical compounds tested across different assays.

PubChem is an open database containing detailed data about chemicals and bioassays, which complements the information from Tox21 and ToxCast.

The Data Collection Plan involves accessing these datasets through Kaggle and other public repositories. Additionally, data preprocessing will be performed to manage missing values, outliers, and duplicate records, ensuring a clean and reliable dataset for analysis.

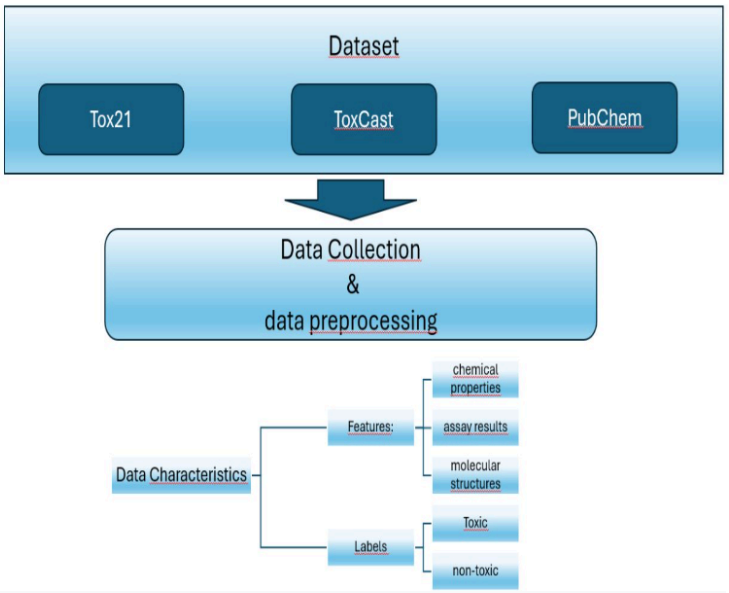
In terms of Data Characteristics, the dataset will consist of:

Features: These include chemical properties, assay results, and molecular structures.

Labels: Toxic versus non-toxic predictions, which will be used as target labels for toxicity classification tasks.

This structured approach will ensure a comprehensive data foundation for further analysis and model building.

Data Gathering Plan



#5 Wireframes for your project (what are all the components of the project, how do they work with each other, what will the end-result look like) - 25 pts