Proposed Title:

EUGTO (pronounced as *"Yugto"*)

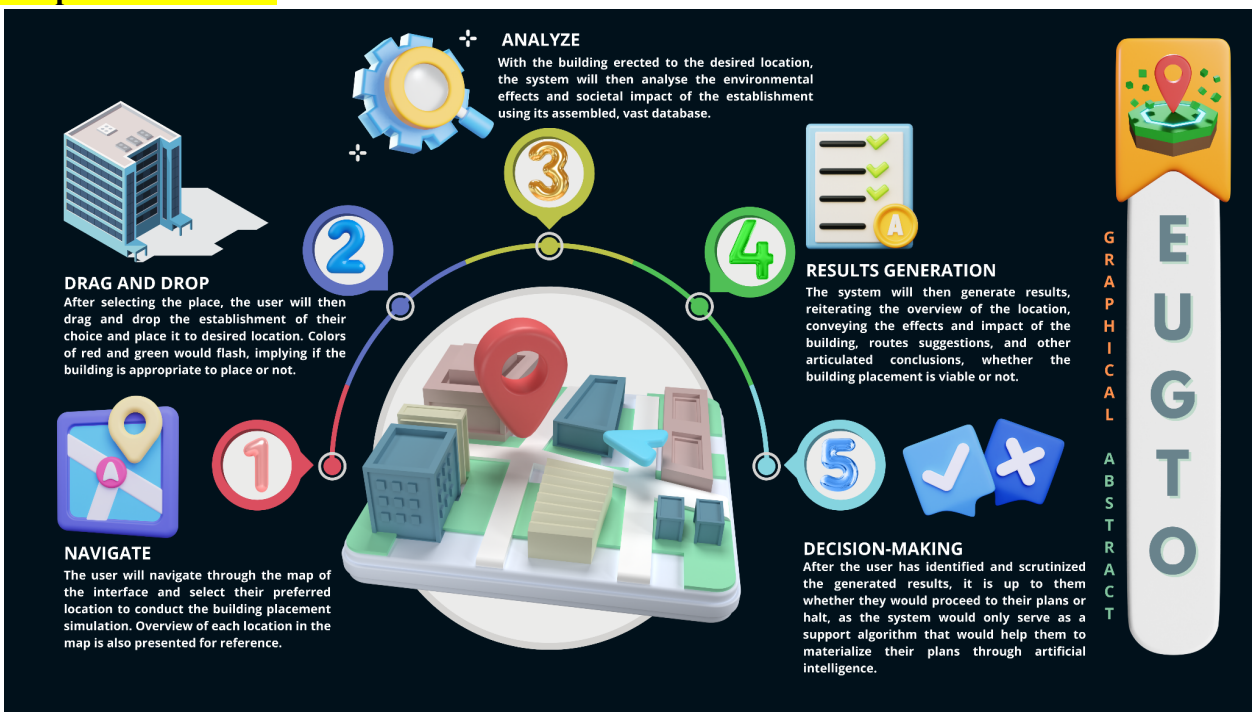
Description:

The proposed system intends to *analyze and forecast the aftermaths* of an establishment by giving it a *"moment"* to be placed in the desired location through the use of artificial intelligence, thus, the name EUGTO.

Keywords:

- Building Placements
- Environmental Impact/Effects
- Suggestive Nature
- Predictions through Moment (Yugto) simulations using Artificial Intelligence
- **E**stablishing **U**rban Building Placements and **G**enerating Environmental & Societal Impact **T**hrough Artificial Intelligence for Smart Cities **O**perations *(Tentative Meaning)*

Graphical Abstract:



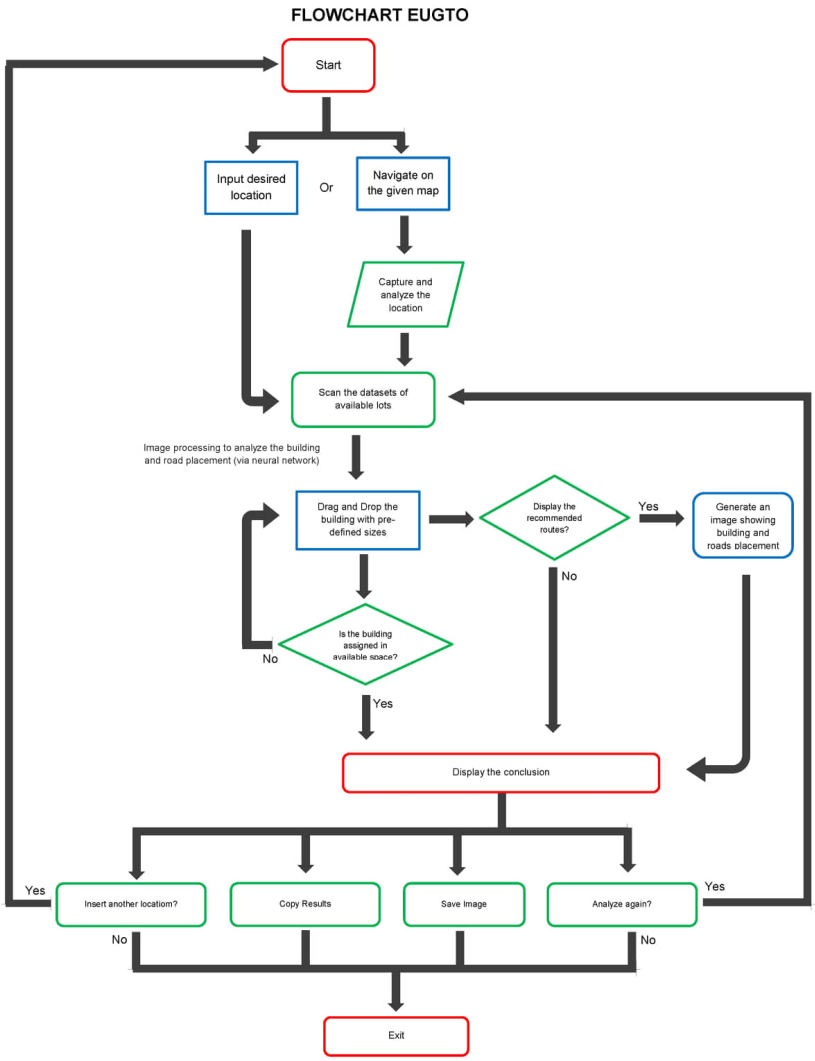
The figure shows the graphical abstract of the A.I project conveying the 5 processes: *Navigate, Drag and Drop, Analyze, Results Generation, and Decision-Making.* The proposed system would have a feature that presents a question, in which the system will bring forth some options to indicate where the user intends to place their establishment, or if they are allowing the system to decide for the most viable location instead.

Upon navigating the Philippine Map, the system would also show a quick overview of the place to help with the selection. After deciding which location to use, the user can then choose from the building templates, varying in sizes and population capacity, and drop them to their selected location, yielding to a simulation in which the system illustrates the condition of the building in that particular place, returning a feedback of routes to be taken or created, attributes of the place with statistics, measure the safety, security, and stability of the establishment in that location, calculating the risks of possible dangers and accidents, either by

natural or artificial disasters, and stating if the establishment is viable to be erected or not. The results will then serve as a supporting document for professionals to help their decision-making process.

For more detailed information, see the flowchart below.

Framework:



Data Acquisition:

This section conveys the viable sources that the team may request to acquire datasets from. It includes government agencies and private associations.

For Available and Alienable and Disposable Areas:
Land Management Bureau (Government)
https://lmb.gov.ph/

Requirement: Create a Request Letter signed by the project adviser and team representative, and send it to the website's contact form in the Contact Us tab.

For Disaster Mapping:
(LiPAD) LiDAR Portal for Archiving and Distribution
https://lipad.dream.upd.edu.ph/?fbclid=IwAR3H9u-XYsmMAB2O4fzuUFIbKK6ip2Jhq183JrHTtALx9BOSd9D8yl9xP_0
https://www.aiddata.org/data/em-dat-phl

For Flood Risk:
ICA Philippines - Flood Risk
https://data.humdata.org/dataset/wfp-geonode-ica-philippines-flood-risk

For Landslide Risk:
NOAH Landslide Hazard Maps
https://asti.dost.gov.ph/coare/data/datasets/noah-landslide-hazard
https://oasishub.co/dataset/philippines-all-provinces-landslide-hazard-maps-noah

For Hazard Risk Mapping:
PHIVOLCS
https://www.phivolcs.dost.gov.ph/index.php/gisweb-hazard-maps

For Earthquake- & Volcano-Related Maps:
PHIVOLCS
https://gisweb.phivolcs.dost.gov.ph/gisweb/earthquake-volcano-related-hazard-gis-information

TIP: Can refer to GeoRiskPH apps like:
https://geoanalytics.georisk.gov.ph/ , https://hazardhunter.georisk.gov.ph/
https://geomapper.georisk.gov.ph/

For Street Views Dataset:
Google Street Views Dataset
https://www.google.com/streetview/
Instant Street View
https://www.instantstreetview.com/ [Extra features require monetary payment]

Open Datasets:
Geospatial Machine Learning for Urban Development [Timestamp: 4:09]
https://youtu.be/DQtV2ikrQxs

Data of Carbon Footprint in different references (Philippines per year):
Worldometer
https://www.worldometers.info/co2-emissions/philippines-co2-emissions/#:~:text=CO2%20emissions%20per%20capita%20in,in%20CO2%20emissions%20per%20capita.
Worldbank

https://data.worldbank.org/indicator/EN.ATM.CO2E.PC?end=2019&locations=PH&start=2019&view=map
Macrotrends
https://www.macrotrends.net/countries/PHL/philippines/carbon-co2-emissions

Data Analysis and Exploration (Algorithms):

This section portrays summaries of each algorithm assigned to every member of the team. The summaries include the definition, the processes, and the examples or relevance of the algorithm to the A.I project at hand. This is a crucial part of the progress since the conclusions articulated here would dictate what processes and platforms the project would center on.

1.) Linear Regression

Definition: It is used to estimate real values (cost of houses, number of calls, total sales etc.) based on continuous variable(s). Here, we establish the relationship between independent and dependent variables by fitting the best line.

How it works:

- Data set can be split into training dataset and testing dataset.
- Training dataset will create mode and can be used to predict the dependent variable of the testing dataset.
- If the accuracy score is not accurate enough and a stronger model wants to be built, the percentage of the datasets allocated to the training and testing datasets can be changed.

Example : Let us say, you ask a child in fifth grade to arrange people in his class by increasing the order of weight, without asking them their weights. He/she would likely look (visually analyze) at the height and build of people and arrange them using a combination of these visible parameters. The child has actually figured out that height and build would be correlated to weight by a relationship.

2.) Logistic Regression:

Definition: It is used to estimate discrete values (Binary values like 0/1, yes/no, true/false) based on a given set of the independent variable(s). In simple words, it predicts the probability of occurrence of an event by fitting data to a logit function. Hence, it is also known as logit regression. Since it predicts the probability, its output values lie between 0 and 1 (as expected).

How it works:

- Analyzes relationships between variables.
- Assign probabilities to discrete outcomes using sigmoid function.
- Converts numerical results into an expression of probability between 0 and 1.0.

Example: To predict whether a political candidate will win or lose an election or whether a high school student will be admitted or not to a particular college. These binary outcomes allow straightforward decisions between two alternatives.

Sources:
[1] https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

[2] https://www.kdnuggets.com/2022/07/logistic-regression-work.html

3.) K means

Definition: categorized as one of the simplest and popular unsupervised ML algorithms. This clustering categorizes the same things altogether, and isolate it to other clusters. K-means looks for a fixed number (k) of clusters in a dataset.

How does it work: Allocation of the centroids (data) → Determine the distances of centroids → Assign the real centroid (middle) of each cluster → centroid repositioning stops. DONE, the algorithm is converged.

- Supports various kinds of distance measures, such as:
    - Euclidean distance measure – If we have a point P and point Q, the euclidean distance is an ordinary straight line
    - Manhattan distance measure - along axes at right angles.
    - A squared euclidean distance measure
    - Cosine distance measure

Example: Academic performance, Diagnostic systems, Search engines, Wireless sensor networks

Programming Language: Python, C++, C, Java, SQL

Steps:

- Data pre-processing
- Finding the optimal number of clusters using the elbow method
- Training the K-Means algorithm on the training data set
- Visualizing the clusters

Source:
https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm#types_of_clustering

4.) Random Forest

Definition: categorized as supervised ML algorithms that is used for classification (handles categorical variables) and regression (handles continuous variables) problems. It consists of many decisions trees. It uses bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree.

- **Two Methods:**
    - Bagging – creates training subset rom sample training data with replacement & the final output is based on **majority voting.**
    - Boosting - It combines weak learners into strong learners by creating sequential models such that the final model has the **highest accuracy**.
- **How does it work:**

Step 1: In Random Forest n number of random records are taken from the data set having k number of records.

Step 2: Individual decision trees are constructed for each sample.

Step 3: Each decision tree will generate an output.

Step 4: Final output is considered based on Majority Voting or Averaging for Classification and regression respectively.

Programming Languages: Python, C++, Java, SQL

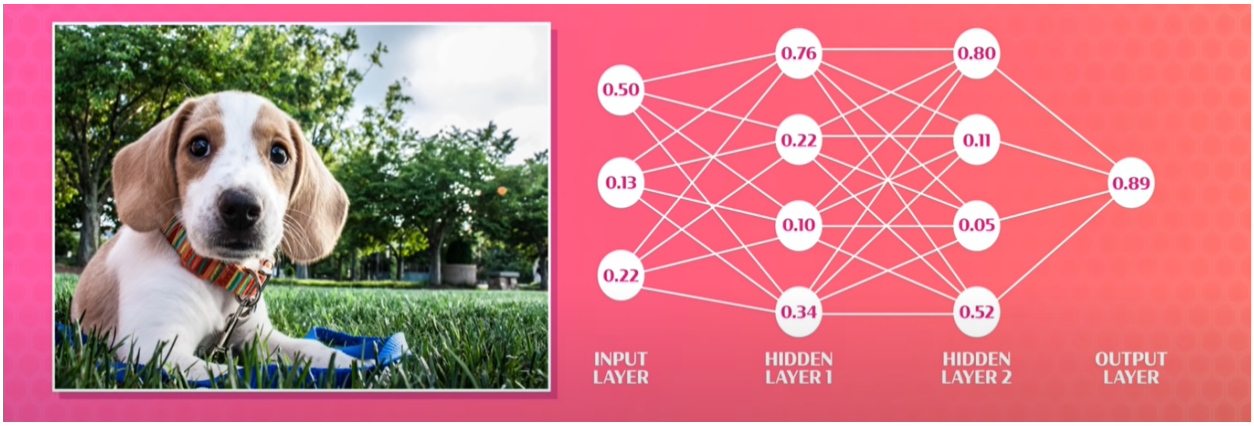
Sources: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/

5.) Neural Network

Definition: Neural Network is a type of an algorithm in artificial intelligence that works similarly to the human brain. It has neurons that are interconnected together, passing information from one another, and divided by layers. T

How it Works: The layers consist of the input layer, in which it detects and receives the information from the external source, then the hidden layers is where most of the detection and analyzing processes take place, then finally, the output layer that will return the results of the system.

Each neuron contains a feature that is being evaluated by the system, and these features are represented by numbers, most commonly known as, “weights”, which makes the system perform necessary recognition and calculations. (See picture below for further illustration).



Sites to be checked for neural network software (Commonly used platform is Python):
[1] https://www.analyticsinsight.net/top-10-must-know-artificial-neural-network-software/
[2]https://realpython.com/python-ai-neural-network/

Related Studies:
[1] https://doi.org/10.1063/1.1144830
[2] https://doi.org/10.1111/0885-9507.00219
[3]https://www2.deloitte.com/ce/en/pages/real-estate/articles/how-ai-can-enhance-urban-planning-asset-management-and-investments.html
[4] https://youtu.be/bUv38kHOHbw