# ArCADia; Investigation of Matching Method for AI Aided CAD Development Prompting Tools

Anushka Polapally

## Introduction

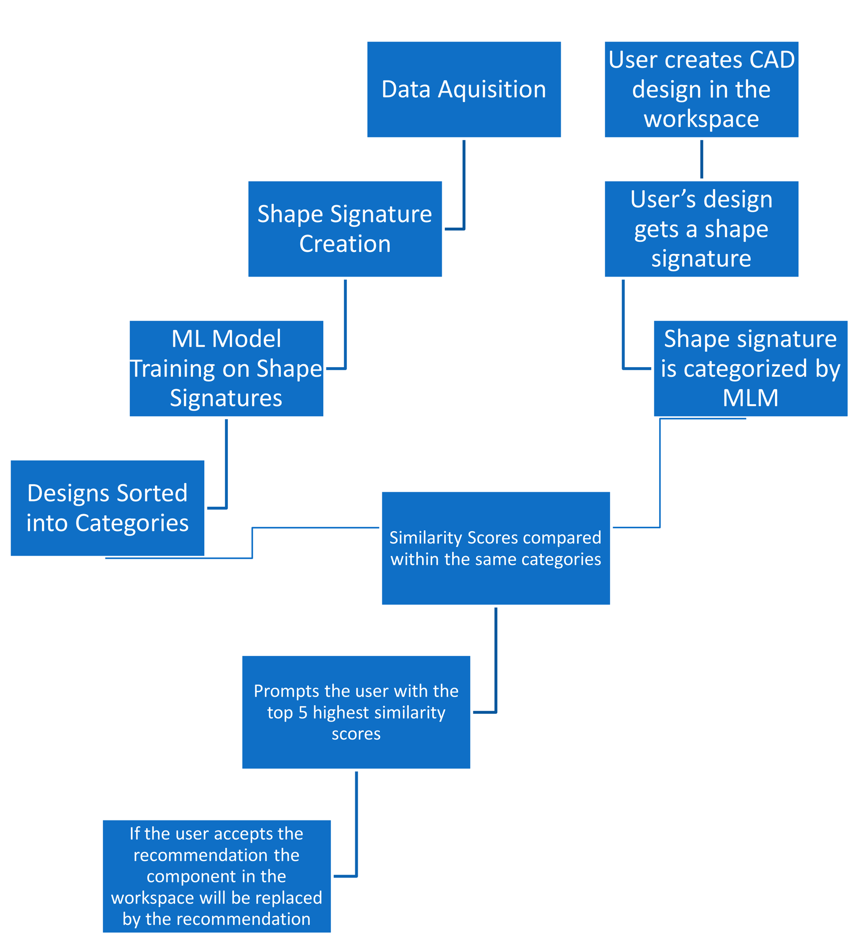
ArCADia is a program designed to give **optimized recommendations** to users during the CAD development process, by utilize a matching process between the user's design and previous designs stored in a database. Using CAD is often a lengthy and tedious process, that requires users to manually recreate components that are already found online, because of the lack of automatic implementation. ArCADia aims to mitigate this issue via, efficient prompts, generated by **similarity matching** with designs already processed.

This project investigates a matching method for user made designs and designs found in the database. The matching method will be validated by using statistical analysis of the similarity found by a ML MATLAB model. This model will utilize various mathematical functions to interpret the geometric data. We will explore the optimization of the proposed algorithm.

## Research Objectives

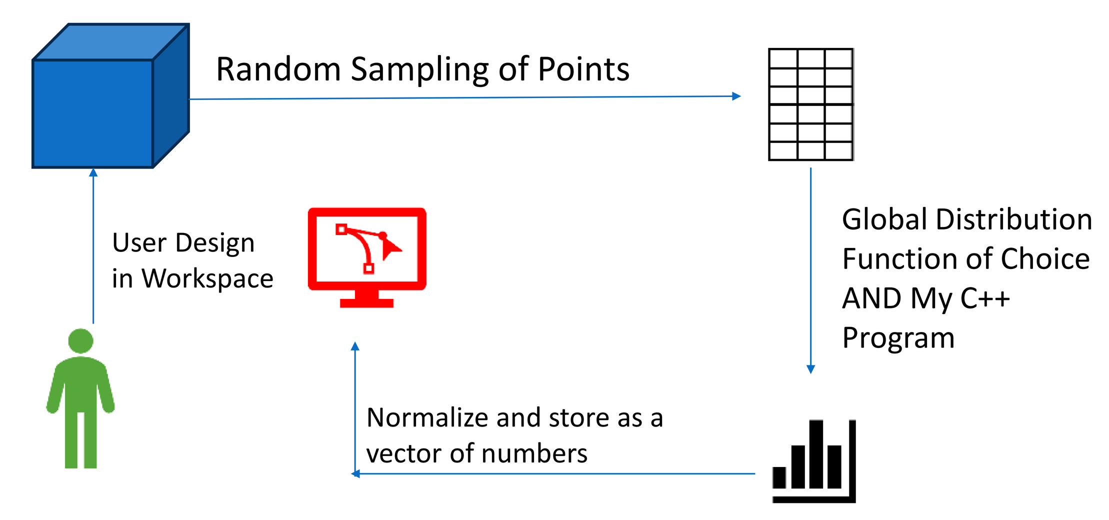
**Problem** - CAD development tools do not offer recommendations that users can implement within their projects to reduce time wasted making preexisting components.
**Research Question** - Can a method be developed that is able to recognize a design that a user is creating, and then recommend a similar refined design from a database for the user to implement?
**Claim** – This research will offer an algorithm that demonstrates the ability to match CAD designs accurately as well as offer an examination into the optimization of this process. It will utilize mathematical functions as well as ML for the key components of the algorithm.
**Primary Endpoint** – The algorithm can perform at >70% accuracy during testing for its top recommendation to the user.

## Algorithm Proposal



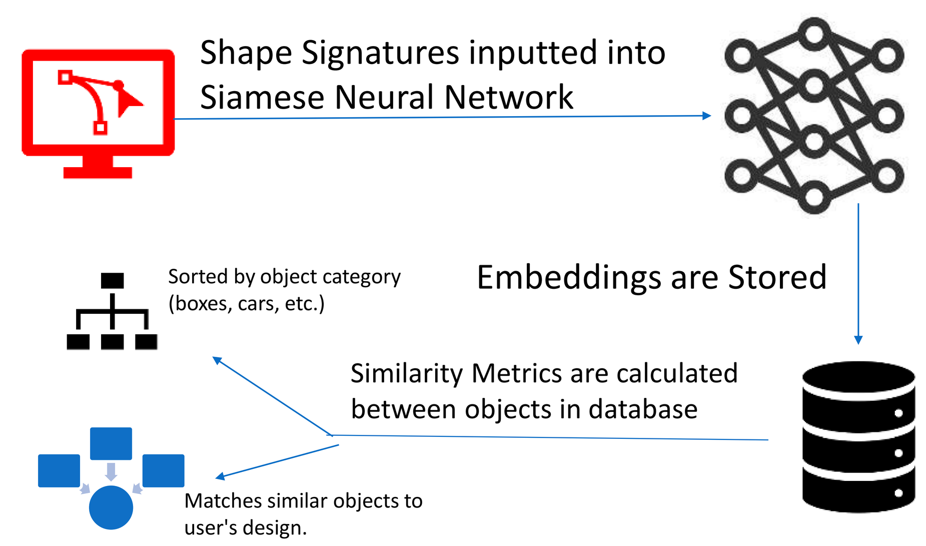

Program Design → Shape Signature Generation → MLM Training → Data Analysis

## Shape Signature Generation




The **shape signature** will represent an object as a shape distribution sampled from a shape function that will measure global properties of the object. These shape signatures were **generated via a histogram, programmed in C++**, that creates the shape signatures with whichever shape distribution function is chosen.

## MLM Training




This project trains a **Deep Metric Learning (DML) model** to match 3D shapes using shape distributions. A **Siamese Neural Network** is trained on pairs of shape signatures, learning to minimize the distance between similar objects and maximize it for dissimilar ones. The model then predicts shape similarities by comparing new shape signatures, enabling fast and robust 3D model retrieval. Each pair of objects will be given a similarity score.

## Optimization

There are 5 options for the global distribution function that can be used:

- A3: Measures the angle between three random points on the surface of a 3D model.
- D1: Measures the distance between a fixed point and one random point on the surface. We use the centroid of the boundary of the model as the fixed point.
- D2: Measures the distance between two random points on the surface.
- D3: Measures the square root of the area of the triangle between three random points on the surface.
- D4: Measures the cube root of the volume of the tetrahedron between four random points on the surface.

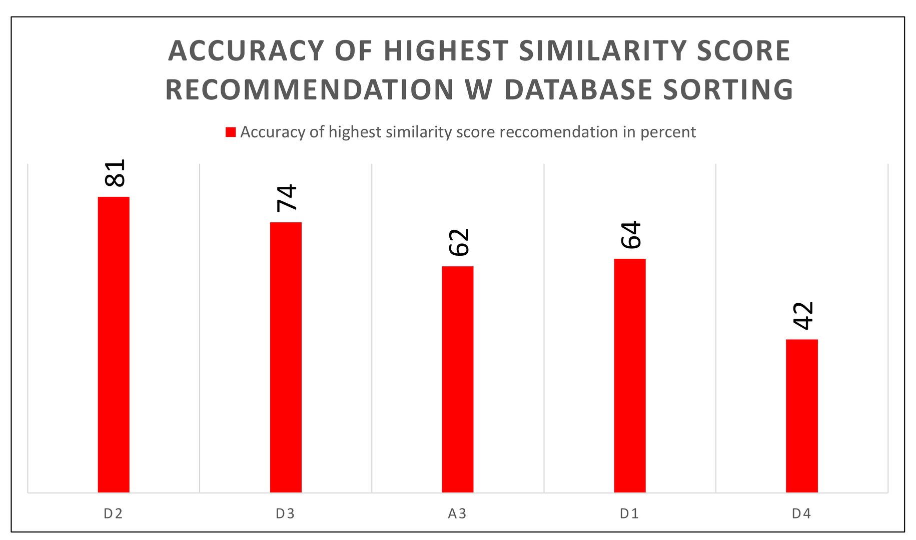
These functions could have different results in accuracy due to varying sensitivity and noise.

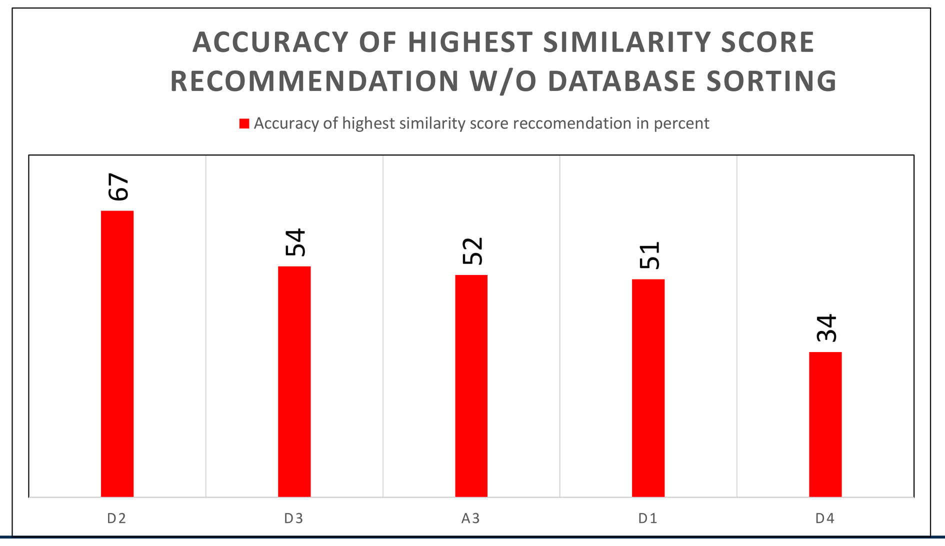
Large variations between designs can be reduced by sorting the initial dataset after MLM training. However, this would require more compute power, so it should be proven necessary.

## Optimization and Examination of Process Using Statistical Analysis

We aim to explore two variabilities: the global distribution function used to create the shape signature, and whether the database is sorted or not before creating a similarity score.

- An initial dataset tagged with the object's category were used. This data was from the ABC database.
- Designs from GrabCAD and other such websites were used to perform as the "user's design"
- Two ML models where trained using the database, one that sorts designs into categories before calculating similarity score w/ the users design, and one that does not.
- The sorted ML model, assigns the user's design a category as well, optimizing search time, but model training is much more compute intensive.
- The design given the highest similarity score to the user's design is checked to see if it has the same category as the user's design.
- This process is repeated for both models, while using each global distribution function.




Optimization Testing Results:

- D2 as the global distribution function generates the most accurate top recommendation.
- Database sorting by the MLM generates significantly more accurate recommendations.

Factor of accuracy D2 vs. other global distribution functions

| | D3 | A3 | D1 | D4 |
|---|---|---|---|---|
| Data Sorted | 1.090 | 1.306 | 1.257 | 1.929 |
| Data Unsorted | 1.241 | 1.288 | 1.314 | 1.971 |

## Conclusion

- The algorithm has high accuracy when using the D2 function and a sorted database by object category.
- By having accurate prompting for the user, it will help **accelerate product design and R&D** by preventing repetitive tedious tasks.
- This tool is parallel to software development tools such as CoPilot.

## Future Work

1. Release a Beta Version of ArCADia
2. Develop a unique global distribution function to increase accuracy of recommendations
3. Increase database size to cover more use cases
4. Expand this algorithm into similar fields that involve ML for geometric data

## Key References

1. Heidari, N., & Iosifidis, A. (2024). Geometric Deep Learning for Computer-Aided Design: A Survey. ArXiv.org. https://arxiv.org/abs/2402.17695
2. Osada, R., Funkhouser, T., Chazelle, B., & Dobkin, D. (n.d.). Matching 3D Models with Shape Distributions. Retrieved January 24, 2025, from https://www.cs.princeton.edu/~funk/smi01.pdf
3. Koch, S., Matveev, A., Jiang, Z., Williams, F., Artemov, A., Burnaev, E., Alexa, M., Zorin, D., & Panozzo, D. (2019, April 30). ABC: A BIG CAD model dataset for Geometric Deep Learning. arXiv.org. https://arxiv.org/abs/1812.06216

All images are created by student unless otherwise specified