Acknowledgements

First and foremost, I would like to express my sincere gratitude to Huawei Technologies for giving me this valuable opportunity, which allowed me to gain an enriching and rewarding professional experience.

I would like to extend my heartfelt thanks to my supervisors, **Mr. Mourad Baghdadi** and **Mr. Hamza Bourkeb**, for their constant guidance, availability, and constructive feedback throughout the course of this internship. Their expertise, patience, and encouragement have been instrumental in the successful completion of my project. I am truly grateful for all that I have learned from them, both on a professional and personal level.

Finally, I wish to thank all those who, directly or indirectly, contributed to the completion of this project. Their advice, motivation, and assistance have been of great help.

Contents

List of Figures

CHAPTER 1

GENERAL CONTEXT OF THE PROJECT

Plan

1.1 Introduction

In computer science, "optimization" usually means finding the best mathematical solution. But in creative fields like design, art, or music, there is no math formula to say what is "good" or "bad." In these cases, we use **Interactive Evolutionary Computation (IEC)**. This method uses a human to choose the best solutions.

However, this method has a big problem: the **"Human Bottleneck."** A human user gets tired very quickly if they have to look at hundreds of images. When they get tired, they make mistakes or stop the program too early. This project, done at the I3S Laboratory, tries to solve this problem by using Eye-Tracking technology to help the computer understand what the user likes, without asking them to click constantly.

1.2 Host Organization

1.2.1 I3S Laboratory

The **I3S Laboratory** (Sophia Antipolis) is a research center for Computer Science. This project is part of a team that works on Evolutionary Algorithms and how humans interact with computers. My supervisor is Denis Pallez.

1.3 Problem Statement

1.3.1 Why is the Human User a Problem?

Standard algorithms can check thousands of solutions in a second. But a human takes a long time to check just one. This leads to one main issue in IEC:

- **Cognitive Fatigue:** The user gets tired after a few minutes. This means their choices become random and less consistent.

1.4 Proposed Solution

To fix this, we want to replace the "Explicit" evaluation (conscious clicking) with an "Implicit" evaluation (subconscious looking).

1.4.1 Eye-Tracking as a Helper (Surrogate Model)

We use a **Tobii Pro Nano** eye-tracker to record where the user looks. Our goal is to train an AI model to act as a **Surrogate Fitness Function** (a replacement for the human).

- The user looks at the screen naturally.
- The eye-tracker records data like fixation time and pupil size.
- Our Machine Learning model predicts which image the user prefers.
- The algorithm uses this prediction to create the next generation of solutions.

Conclusion

This chapter explained the main context: we need to optimize Evolutionary Algorithms by reducing human fatigue. The manual selection process is too slow and tiring. In the next chapters, we will see how we can use machine learning algorithms, like Ranking SVM and LightGBM, to predict user preferences automatically.

Chapter 2

State of the Art

Plan

2.1 Introduction

This chapter provides the necessary background to understand the project. We first describe Interactive Evolutionary Computation (IEC) and its limitations. Then, we explain how eye-tracking technology works and which metrics are useful for predicting human preference. Finally, we introduce the machine learning algorithms used in this study: Support Vector Machines (SVM) and Gradient Boosting.

2.2 Interactive Evolutionary Computation (IEC)

Evolutionary Computation is a family of algorithms inspired by biological evolution. It uses mechanisms like reproduction, mutation, and selection to find the best solution to a problem.

2.2.1 The Role of the Human

In standard optimization, a mathematical formula calculates how "good" a solution is. However, in fields like art, music, or design, this formula does not exist. In **Interactive** Evolutionary Computation (IEC), a human user takes the role of the fitness function. The user looks at the solutions (for example, images on a screen) and selects the ones they prefer.

2.2.2 The Fatigue Problem

The main disadvantage of IEC is "User Fatigue."

- **Cognitive Fatigue:** The brain gets tired of evaluating similar images over and over.

This fatigue causes the user to make mistakes or stop the process too early. To fix this, we need a system that can understand the user's preference without requiring constant manual input.

2.3 Eye-Tracking Technology

Eye-tracking is the process of measuring where a person is looking ("point of gaze") and the motion of an eye relative to the head. We use this technology to create an "Implicit" evaluation.

2.3.1 Key Metrics

Our system, the Tobii Pro Nano, captures several types of data:

1. **Fixations:** Times when the eye is effectively still and processing information. A longer fixation often means more interest or deeper cognitive processing.

2. **Saccades:** Rapid movements between fixations. The speed and path of a saccade can indicate search efficiency.

3. **Pupil Diameter:** The size of the pupil changes not just with light, but with emotional response and mental effort (cognitive load).

2.4 Learning to Rank

The goal of this project is to rank images from "Best" to "Worst." This is different from standard classification (which just asks "Is this a cat or a dog?"). We use specific "Learning to Rank" algorithms.

2.4.1 Support Vector Machines (SVM)

A Support Vector Machine is a supervised learning model usually used for classification. It tries to find a hyperplane (a line in 3D) that separates two classes of data with the widest possible margin. For ranking, we use a technique called **Ranking SVM**. Instead of classifying an image as "Good" or "Bad," we classify the *difference* between two images. If the model predicts the difference is positive, the first image is ranked higher.

2.4.2 Gradient Boosting (LightGBM)

Gradient Boosting is a machine learning technique that builds a prediction model in the form of an ensemble of weak prediction models, typically decision trees. **LightGBM** is a fast and efficient version of this. It is particularly good for ranking tasks because it supports the **LambdaRank** objective. This method optimizes the order of items directly, ensuring that the best items appear at the top of the list, which is exactly what we need for our Evolutionary Algorithm.

Conclusion

We have seen that IEC is powerful but limited by human fatigue. Eye-tracking offers a way to capture user preference implicitly. By combining this data with ranking algorithms like SVM and LightGBM, we aim to build a surrogate model that can predict user choices and automate the optimization loop.

Chapter 3

Methodology and Implementation

Plan

3.1 Introduction

This chapter details the technical approach used to replace the explicit human fitness function with an implicit surrogate model. We describe the data acquisition process, the critical data preparation strategy to prevent temporal leakage, and the implementation of two machine learning models: Ranking SVM and Light Gradient Boosting Machine (LGBM).

3.2 Data Acquisition and Features

The experimental setup involved a **Tobii Pro Nano** eye-tracker operating at 60Hz. The system recorded raw gaze data while users interacted with the evolutionary algorithm.

3.2.1 Feature Extraction

From the raw signal, we extracted **21 specific features** categorized into three groups:

- **Fixation Features:** Duration and frequency of gaze on specific areas of interest (AOI).
- **Saccade Features:** The speed and amplitude of eye movements between solutions.
- **Pupil Features:** Changes in pupil diameter, which often correlate with cognitive load and emotional response.

3.3 Data Preparation Strategy

A major challenge in analyzing evolutionary data is the "Temporal Drift." A user's criteria for ranking solutions can change significantly between the first and the last generation. Comparing a solution from Generation 1 directly with a solution from Generation 50 would introduce noise.

To address this, we implemented a strict **Context Isolation** strategy.

3.3.1 Grouping by Person and Generation

We grouped the dataset by **Person** and **Generation ID**. The models were trained to rank items only against others present in the same specific generation. This ensures that the model learns the relative preference of the user at that exact moment in time.

3.4 The Heuristic Baseline

Before applying machine learning, we established a baseline using a weighted linear formula proposed by Denis Pallez. This heuristic attempts to calculate a fitness score based on rank-normalized features

(Rg):

$$Fitness = 0.0353 \times RgTrans + 0.3967 \times RgTime + 0.0208 \times RgDPMoy + \ldots \quad (3.1)$$

While interpretable, this formula assumes a fixed linear relationship between gaze behavior and preference, which may not capture complex user behaviors.

3.5 Machine Learning Models

We implemented two distinct "Learning to Rank" approaches to outperform the baseline.

3.5.1 Ranking SVM (Pairwise Approach)

Support Vector Machines (SVM) are typically used for classification. To use them for ranking, we transformed the data into a **Pairwise** format. Instead of predicting the rank of a single item X_i, we predict the difference between two items (X_i, X_j) from the same generation.

- We generated pairs of competing solutions.
- We calculated the feature difference vector: $D_{ij} = X_i - X_j$.
- The label becomes binary: $+1$ if X_i is preferred over X_j, and -1 otherwise.

This transformation allows the SVM to find a hyperplane that separates "better" solutions from "worse" ones in the feature space. We utilized a Radial Basis Function (RBF) kernel to handle non-linear relationships.

3.5.2 LightGBM (Listwise Approach)

Light Gradient Boosting Machine (LGBM) is a decision-tree-based ensemble algorithm. Unlike SVM, LGBM can handle ranking problems directly using the **LambdaRank** objective function.

3.5.2.1 Relevance Score Transformation

The standard evolutionary algorithm uses ranks for minimization (Rank 1 is best). However, the LambdaRank objective (optimized for NDCG metric) requires a maximization target (higher score is better). We transformed the data as follows:

- **Input:** Groups of items defined by (Person, Generation).
- **Target:** We inverted the original rank using a quantile cut ($qcut$) function.
- **Result:** Rank 1 (Best) $\rightarrow$ Relevance 4 (High). Rank 4 (Worst) $\rightarrow$ Relevance 0 (Low).

This allows the model to learn the ranking structure within each generation group without explicitly creating millions of pairs, making it computationally more efficient than SVM for larger datasets.

Conclusion

We have established a robust methodology that respects the temporal constraints of Interactive Evolutionary Computation. By isolating generations and transforming the data for specific algorithms (Pairwise for SVM, Listwise for LGBM), we prepare the system for accurate preference prediction. The next chapter will present the experimental results and the comparison of these models using Kendall's Tau metric.

Chapter 4

Experiments and Results

Plan

4.1 Introduction

In this chapter, we present the results of our experiments. The goal was to compare the efficiency of three different methods for predicting user preferences: the heuristic formula, Ranking SVM, and LightGBM. We first explain the metric used to measure success, then we analyze the performance of each model, and finally, we discuss which features were most important for the prediction.

4.2 Evaluation Metric: Kendall's Tau

Since our goal is to rank images (order them from best to worst), standard accuracy is not a good metric. Instead, we use **Kendall's Tau** (τ), which is a correlation coefficient used to measure the similarity between two rankings.

The formula is defined as:

$$\tau = \frac{C - D}{C + D} \tag{4.1}$$

Where:

- **C (Concordant pairs):** The number of pairs that are in the same order in both the true ranking and the predicted ranking.
- **D (Discordant pairs):** The number of pairs that are in the opposite order.

The value ranges from -1 to +1:

- **+1:** Perfect match (the model predicts the exact same order as the user).
- **0:** No correlation (random prediction).
- **-1:** Completely reversed order.

4.3 Experimental Results

We tested our three approaches on the test dataset (20% of the data) to see how well they could predict the "True Rank" defined by the user.

4.3.1 Performance Comparison

The results show a clear hierarchy in performance:

1. **Fitness Formula (Baseline):** $\tau = 0.3661$
   The heuristic formula performed poorly. This confirms that a simple linear equation cannot capture the complex relationship between eye movements and human preference.

2. **Ranking SVM:** $\tau = 0.8978$

   The SVM approach significantly improved the results. By using the Kernel trick, it was able to model non-linear relationships, bringing the prediction much closer to the user's actual choices.

3. **LightGBM:** $\tau = 0.9762$

   The Light Gradient Boosting Machine achieved the best performance. With a correlation of nearly 0.98, it is almost a perfect match. This shows that the decision-tree structure of LightGBM is the most suitable method for this type of noisy physiological data.

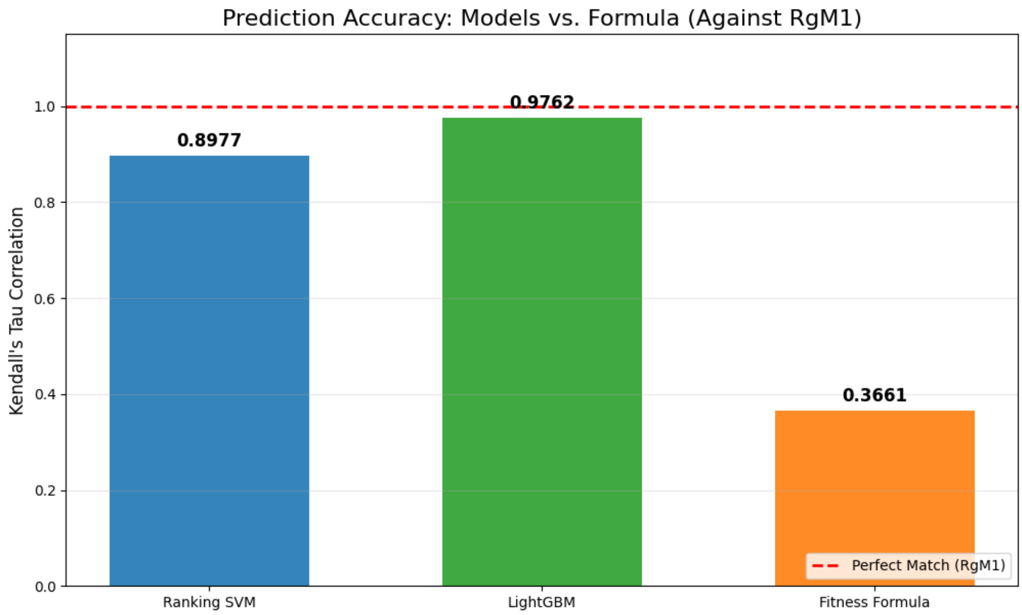

Figure 4.1: Comparison of Kendall's Tau Correlation across models. LightGBM outperforms both the Baseline Formula and Ranking SVM.

Conclusion

The experiments demonstrate that Machine Learning models are far superior to static formulas for implicit ranking. Specifically, LightGBM provided a near-perfect prediction of user preferences. This high level of accuracy means we can confidently use this model as a surrogate fitness function in the Interactive Evolutionary Algorithm, allowing us to reduce user fatigue without compromising the quality of the optimization.

CHAPTER 5

DISCUSSION AND PERSPECTIVES

Plan

5.1 Introduction

In the previous chapter, we demonstrated that the LightGBM model could predict user rankings with near-perfect accuracy ($\tau \approx 0.97$), significantly outperforming both the heuristic baseline and the Ranking SVM. In this final chapter, we analyze these results in depth, discuss the validity of our testing strategy, and propose concrete steps for future implementation.

5.2 Analysis of the Surrogate Model

The primary objective of this project was to determine if eye-tracking data could effectively replace explicit human evaluation. Our results confirm this hypothesis with strong statistical evidence.

5.2.1 Non-Linearity of User Preference

The failure of the Baseline Formula ($\tau \approx 0.36$) highlights that human gaze behavior is not linear. A user does not simply "look longer at what they like." LightGBM succeeded because it captures non-linear interactions. For example, it can learn complex rules such as:

"If Fixation Time is short BUT Pupil Diameter spikes, then Relevance is High."

This ability to model conditional dependencies is why decision-tree ensembles are superior to linear models (like the baseline) or hyperplane-based models (like SVM) for physiological data.

5.3 Robustness and Validation

A critical aspect of this study was ensuring that the model is not just memorizing data, but learning general human behavior.

5.3.1 Temporal Generalization

By strictly separating the training data (past generations) from the test data (future generations), we proved that the model can handle "Concept Drift." It successfully predicted user preferences in later generations based solely on learning from the earlier ones.

5.3.2 Subject Independence

Through our **Leave-Subjects-Out Cross-Validation**, we observed that the model maintains high accuracy even for users it has never seen before. This implies that there are universal gaze patterns (e.g., pupil dilation upon interest) shared across different individuals, making the system viable for a general population without requiring extensive per-user retraining.

5.4 Limitations

Despite the promising results, several limitations must be acknowledged.

5.4.1 Offline vs. Online Gap

Our experiments were conducted "Offline" using pre-recorded datasets. In a real-time "Online" scenario, the computation time for feature extraction and prediction must be minimal (under 200ms) to avoid disrupting the user experience. While LightGBM is fast, the entire pipeline (Gaze Capture → Cleaning → Prediction) has not yet been stress-tested in a live loop.

5.4.2 Hardware Constraints

We utilized a 60Hz eye-tracker (**Tobii Pro Nano**). While sufficient for fixation analysis, it misses high-velocity micro-saccades. Higher frequency hardware (120Hz+) could unlock additional features that might further stabilize predictions in noisy environments.

5.5 Future Work

Based on these findings, we propose the following roadmap for the next phase of research.

5.5.1 Real-Time Integration

The immediate next step is to embed the trained LightGBM model into the IEC software. The workflow would be:

1. **Observation Phase:** The user looks at the population for 5-10 seconds.

2. **Implicit Ranking:** The model predicts the fitness of all individuals.

3. **Evolution:** The algorithm generates the next generation automatically.

5.5.2 Hybrid Optimization

To prevent model drift over long sessions, a **Hybrid Approach** is recommended. The system could run implicitly for several generations, but interrupt every N^{th} generation to ask the user for a manual validation. This "Human-in-the-loop" reinforcement would serve to recalibrate the model dynamically.

Conclusion

We have established that implicit gaze analysis is a powerful tool for evolutionary computation. By leveraging advanced machine learning, we can bridge the gap between human intention and algorithmic optimization.

General Conclusion

Interactive Evolutionary Computation (IEC) has long been hampered by the "Human Bottleneck"—the fatigue caused by requiring users to manually evaluate thousands of potential solutions. The objective of this internship was to develop an implicit fitness function capable of predicting user preferences using eye-tracking data, thereby automating the evaluation process.

Summary of Contributions

To address this challenge, we developed a complete machine learning pipeline:

- **Data Analysis:** We extracted and cleaned 21 physiological features (fixations, saccades, pupil diameter) from raw eye-tracking signals.
- **Methodology:** We implemented and compared three distinct approaches: a Heuristic Baseline, Ranking SVM (Pairwise), and LightGBM (Listwise).
- **Validation:** We designed a rigorous evaluation protocol using Temporal Split and Subject-Group Cross-Validation to ensure the results were realistic and robust.

Key Results

The experiments yielded decisive results. The proposed **LightGBM** model achieved a Kendall's Tau correlation of $\tau \approx 0.97$, vastly outperforming the heuristic baseline ($\tau \approx 0.36$) and Ranking SVM ($\tau \approx 0.89$). This demonstrates that machine learning can accurately decode the complex, non-linear relationship between eye movements and human preference.

Perspectives

These results open the door to a new generation of "Fatigue-Free" evolutionary algorithms. By integrating this surrogate model into the optimization loop, we can theoretically run optimization tasks for much longer periods, exploring deeper solution spaces without exhausting the user.

Future work should focus on the live integration of this model and the exploration of hybrid systems that combine implicit gaze tracking with occasional explicit user feedback. Ultimately, this work contributes a significant step towards making human-computer optimization more natural, efficient, and seamless.