**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC ĐẠI NAM**

Logo, company name

Description automatically generated

**BÀI TẬP LỚN**

**TÊN HỌC PHẦN: TRÍ TUỆ NHÂN TẠO**

**Giáo viên hướng dẫn: TS. Trần Đăng Công**

**Sinh viên thực hiện:**

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Mã sv** | **Họ và tên** | **Lớp** |
| **1** | **1771020660** | **Đinh Văn Thường** | **CNTT 17-08** |
| **2** | **1771020572** | **Nguyễn Đình Quang** | **CNTT 17-08** |
| **3** | **1771020769** | **Đàm Minh Vương** | **CNTT 17-08** |

**Hà Nội, năm 2025**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC ĐẠI NAM**

Logo, company name

Description automatically generated

**BÀI TẬP LỚN**

**TÊN HỌC PHẦN: TRÍ TUỆ NHÂN TẠO**

**ĐỀ TÀI: ỨNG DỤNG GIẢI THUẬT DI TRUYỀN DỰ ĐOÁN THỊ TRƯỜNG TÀI CHÍNH**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **STT** | **Mã Sinh Viên** | **Họ và Tên** | **Ngày Sinh** | **Điểm** | |
| **Bằng Số** | **Bằng Chữ** |
| **1** | **1771020660** | **Đinh Văn Thường** | **10/01/2005** |  |  |
| **2** | **1771020572** | **Nguyễn Đình Quang** | **10/06/2005** |  |  |
| **3** | **1771020769** | **Đàm Minh Vương** | **20/06/2005** |  |  |

**CÁN BỘ CHẤM THI**

**Hà Nội, năm 2025**

**LỜI NÓI ĐẦU**

AI, còn được biết đến như Trí tuệ nhân tạo, là một công nghệ có khả năng giải quyết vấn đề như con người. Cách thức hoạt động của AI dường như mô phỏng trí tuệ của con người – nó có thể nhận dạng hình ảnh, làm thơ và đưa ra dự đoán dựa trên dữ liệu.

Các tổ chức hiện đại thu thập khối lượng dữ liệu cực lớn từ nhiều nguồn khác nhau như cảm biến thông minh, nội dung do con người tạo ra, công cụ giám sát và bản ghi hệ thống. Công nghệ trí tuệ nhân tạo phân tích và sử dụng dữ liệu để hỗ trợ hoạt động kinh doanh một cách hiệu quả. Ví dụ: công nghệ AI có thể phản hồi các cuộc trò chuyện của con người trong lĩnh vực hỗ trợ khách hàng, tạo hình ảnh và văn bản gốc để tiếp thị và đưa ra các đề xuất thông minh để phân tích.

Cuối cùng, mục tiêu của trí tuệ nhân tạo là làm cho phần mềm thông minh hơn để đáp ứng tương tác người dùng tùy chỉnh và giải quyết vấn đề phức tạp.

MỤC LỤC

**[LỜI NÓI ĐẦU](#_Toc193428442)** [3](#_Toc193428442)

**[CHƯƠNG 1: KHÁI NIỆMVỀ TRÍ TUỆ NHÂN TẠO (AI)](#_Toc193428443)** [5](#_Toc193428443)

**[1.1.](#_Toc193428444)****[Khái niệm về Trí Tuệ Nhân Tạo (AI)](#_Toc193428444)** [5](#_Toc193428444)

**[1.2.](#_Toc193428445)****[Lịch sử phát triển của AI](#_Toc193428445)** [5](#_Toc193428445)

**[1.3 AI trong tương lai](#_Toc193428446)** [5](#_Toc193428446)

**[1.4. AI trong doanh nghiệp](#_Toc193428447)** [6](#_Toc193428447)

**[1.5 AI hoạt động như thế nào](#_Toc193428448)** [6](#_Toc193428448)

**[1.6. Các lĩnh vực ứng dụng AI](#_Toc193428449)** [AI đã và đang được ứng dụng rộng rãi trong nhiều lĩnh vực: 6](#_Toc193428449)

**[1.7. Thách thức và định hướng phát triển AI](#_Toc193428450)** [Mặc dù AI mang lại nhiều lợi ích, nó vẫn tồn tại những thách thức như: 7](#_Toc193428450)

**[CHƯƠNG 2: THỰC HÀNH VỀ ĐỀ TÀI](#_Toc193428451)** [8](#_Toc193428451)

**[2.0. Bối cảnh & Tầm quan trọng](#_Toc193428452)** [8](#_Toc193428452)

***[2.0.1. Mục tiêu của đề tài](#_Toc193428453)*** [8](#_Toc193428453)

**[2.1. Phân tích bài toán.](#_Toc193428454)** [8](#_Toc193428454)

***[2.1.1. Đầu vào của hệ thống](#_Toc193428455)*** [8](#_Toc193428455)

[2.](#_Toc193428456) **[Các chỉ số kỹ thuật phổ biến trong giao dịch](#_Toc193428456)**[: 8](#_Toc193428456)

[3.](#_Toc193428457) **[Khoảng thời gian dự đoán](#_Toc193428457)**[: 9](#_Toc193428457)

***[2.1.2. Đầu ra của hệ thống](#_Toc193428458)*** [9](#_Toc193428458)

**[2.2. Ứng dụng Giải thuật Di truyền (Genetic Algorithm - GA) trong bài toán](#_Toc193428459)** [9](#_Toc193428459)

***[2.2.1. Giải thuật di truyền là gì?](#_Toc193428460)*** [9](#_Toc193428460)

***[2.2.2. Ứng dụng GA trong dự đoán giá cổ phiếu](#_Toc193428461)*** [10](#_Toc193428461)

**[2.3 Phân tích code](#_Toc193428462)** [10](#_Toc193428462)

**[2.4 Chạy code và phân tích kết quả](#_Toc193428463)** [17](#_Toc193428463)

**[CHƯƠNG 3: CÁC ĐẶC ĐIỂM CỦA THUẬT TOÁN](#_Toc193428464)** [20](#_Toc193428464)

**[3.1 Giới thiệu về thuật toán di truyền](#_Toc193428465)** [20](#_Toc193428465)

***[3.1.1 Khái niệm thuật toán di truyền](#_Toc193428466)*** [20](#_Toc193428466)

***[3.1.2 Các bước cơ bản của thuật toán di truyền](#_Toc193428467)*** [20](#_Toc193428467)

***[3.1.3 Ứng dụng của thuật toán di truyền](#_Toc193428468)*** [20](#_Toc193428468)

**[3.2 Cấu trúc của thuật toán di truyền](#_Toc193428469)** [21](#_Toc193428469)

***[3.2.1 Khởi tạo quần thể (Population Initialization)](#_Toc193428470)*** [21](#_Toc193428470)

***[3.2.2 Hàm đánh giá độ thích nghi (Fitness Function)](#_Toc193428471)*** [21](#_Toc193428471)

***[3.2.3 Chọn lọc tự nhiên (Selection)](#_Toc193428472)*** [22](#_Toc193428472)

***[3.2.4 Lai ghép (Crossover)](#_Toc193428473)*** [22](#_Toc193428473)

***[3.2.5 Đột biến (Mutation)](#_Toc193428474)*** [22](#_Toc193428474)

***[3.2.6 Đột biến (Mutation)](#_Toc193428475)*** [23](#_Toc193428475)

**[3.3 Ứng dụng thuật toán di truyền trong giao dịch chứng khoán](#_Toc193428476)** [23](#_Toc193428476)

**[3.4 Cải tiến thuật toán di truyền](#_Toc193428477)** [23](#_Toc193428477)

**[CHƯƠNG 4: ỨNG DỤNG THUẬT TOÁN DI TRUYỀN TRONG GIAO DỊCH CHỨNG KHOÁN](#_Toc193428478)** [24](#_Toc193428478)

**[4.1 Giới thiệu](#_Toc193428479)** [24](#_Toc193428479)

**[4.2 Mô hình giao dịch dựa trên thuật toán di truyền](#_Toc193428480)** [24](#_Toc193428480)

***[4.2.1 Thu thập dữ liệu lịch sử](#_Toc193428481)*** [24](#_Toc193428481)

***[4.2.2 Xây dựng tập hợp chiến lược ban đầu](#_Toc193428482)*** [24](#_Toc193428482)

***[4.2.3 Đánh giá chiến lược bằng Backtesting](#_Toc193428483)*** [24](#_Toc193428483)

***[4.2.4 Tiến hóa quần thể chiến lược](#_Toc193428484)*** [25](#_Toc193428484)

*[4.2.5 Kiểm tra hiệu suất chiến lược tối ưu](#_Toc193428485)* [25](#_Toc193428485)

**[4.3 Đánh giá hiệu suất thuật toán](#_Toc193428486)** [25](#_Toc193428486)

**[4.4 Những hạn chế của thuật toán di truyền trong giao dịch chứng khoán](#_Toc193428487)** [26](#_Toc193428487)

**[4.5 Hướng phát triển trong tương lai](#_Toc193428488)** [26](#_Toc193428488)

**[KẾT LUẬN](#_Toc193428489)** [27](#_Toc193428489)

**[DANH MỤC TÀI LIỆU THAM KHẢO](#_Toc193428490)** [31](#_Toc193428490)

# **CHƯƠNG 1: KHÁI NIỆMVỀ TRÍ TUỆ NHÂN TẠO (AI)**

* 1. **Khái niệm về Trí Tuệ Nhân Tạo (AI)**

Trí tuệ nhân tạo (Artificial Intelligence - AI) là một nhánh của khoa học máy tính nhằm tạo ra những hệ thống có khả năng thực hiện các nhiệm vụ thông minh tương tự như con người. AI bao gồm nhiều lĩnh vực như học máy (Machine Learning), xử lý ngôn ngữ tự nhiên (NLP), thị giác máy (Computer Vision), v.v.

* 1. **Lịch sử phát triển của AI**

Năm 1943, Warren McCulloch và Walter Pitts đã đề xuất một mô hình tế bào thần kinh nhân tạo, đặt nền móng cho mạng nơ-ron, công nghệ cốt lõi của AI.  
  
Không lâu sau đó, vào năm 1950, Alan Turing đã xuất bản cuốn “Máy tính và trí tuệ”, giới thiệu khái niệm về Phép thử Turing để đánh giá trí thông minh của máy tính.  
  
Điều này dẫn đến việc các sinh viên tốt nghiệp Marvin Minsky và Dean Edmonds xây dựng cỗ máy mạng thần kinh đầu tiên với tên gọi SNARC, Frank Rosenblatt phát triển Perceptron, một trong những mô hình sớm nhất của mạng nơ-ron, và Joseph Weizenbaum tạo ra ELIZA, một trong những chatbot đầu tiên mô phỏng phương pháp tâm lý trị liệu Rogerian từ năm 1951 đến năm 1969.

Từ năm 1969 đến năm 1979 Marvin Minsky đã chứng minh những hạn chế của mạng nơ-ron, gây ra sự suy giảm tạm thời trong việc nghiên cứu mạng nơ-ron. “Mùa đông AI” đầu tiên xảy ra do giảm kinh phí và hạn chế về phần cứng và máy tính.

**1.3 AI trong tương lai**

Các công nghệ trí tuệ nhân tạo hiện tại đều hoạt động trong một tập hợp các thông số đã xác định trước. Ví dụ: các mô hình AI được đào tạo về nhận dạng và tạo hình ảnh không thể xây dựng trang web.

Trí tuệ nhân tạo tổng quát (AGI) là một lĩnh vực nghiên cứu AI lý thuyết nhằm cố gắng tạo ra phần mềm có trí thông minh giống con người và có khả năng tự học. Mục tiêu là để phần mềm thực hiện các nhiệm vụ mà nó không nhất thiết phải được đào tạo hoặc phát triển cho những công việc đó.

AGI là một mục tiêu theo đuổi lý thuyết để phát triển các hệ thống AI có khả năng tự kiểm soát tự chủ, tự hiểu biết hợp lý và khả năng học các kỹ năng mới. Nó có thể giải quyết các vấn đề phức tạp trong bối cảnh và tình huống mà nó không được dạy trong quá trình tạo ra. AGI với khả năng của con người vẫn là một khái niệm lý thuyết và mục tiêu nghiên cứu. Đó là một tương lai tiềm năng của AI.

**1.4. AI trong doanh nghiệp**

Chatbot và trợ lý thông minh dựa trên AI tham gia vào các cuộc trò chuyện phức tạp hơn và giống con người hơn. Chúng có thể hiểu bối cảnh và tạo ra các câu trả lời mạch lạc cho ngôn ngữ tự nhiên phức tạp và các truy vấn của khách hàng. Chúng có khả năng xuất sắc trong hỗ trợ khách hàng, trợ giúp ảo và tạo nội dung để cung cấp các tương tác được cá nhân hóa. Khả năng học tập liên tục của các mô hình này cho phép chúng thích ứng và cải thiện hiệu năng theo thời gian, nâng cao trải nghiệm và hiệu quả của người dùng.

Ví dụ: [Deriv](https://aws.amazon.com/solutions/case-studies/deriv-case-study/), một trong những nhà môi giới trực tuyến lớn nhất thế giới, phải đối mặt với những thách thức khi truy cập một lượng lớn dữ liệu được phân phối trên nhiều nền tảng khác nhau. Công ty đã triển khai một trợ lý dựa trên AI để truy xuất và xử lý dữ liệu từ nhiều nguồn trong hoạt động hỗ trợ khách hàng, tiếp thị và tuyển dụng. Nhờ có AI, Deriv đã giảm 45% thời gian đào tạo nhân viên mới và giảm 50% thời gian tuyển dụng.

**1.5 AI hoạt động như thế nào**

Hệ thống trí tuệ nhân tạo sử dụng một loạt các công nghệ để hoạt động. Các chi tiết cụ thể sẽ khác nhau, nhưng các nguyên tắc cốt lõi vẫn giữ nguyên: chúng chuyển đổi tất cả các loại dữ liệu, chẳng hạn như văn bản, hình ảnh, video và âm thanh, thành các biểu diễn dạng số và xác định theo toán học các kiểu mẫu và mối quan hệ giữa chúng. Do đó, các công nghệ trí tuệ nhân tạo đòi hỏi phải được đào tạo – chúng được tiếp xúc với khối lượng lớn các tập dữ liệu hiện có để "học" – tương tự như con người học từ các kho lưu trữ kiến thức hiện có. Dưới đây là một số công nghệ làm cho trí tuệ nhân tạo hoạt động.

**1.6. Các lĩnh vực ứng dụng AI**  
AI đã và đang được ứng dụng rộng rãi trong nhiều lĩnh vực:

* Y tế: Chẩn đoán bệnh, phát hiện ung thư, robot phẫu thuật.
* Giao thông: Xe tự lái, hệ thống dẫn đường thông minh.
* Tài chính: Dự báo thị trường, phát hiện gian lận.
* Giáo dục: Gia sư ảo, hệ thống học tập thông minh.

**1.7. Thách thức và định hướng phát triển AI**  
Mặc dù AI mang lại nhiều lợi ích, nó vẫn tồn tại những thách thức như:

* Vấn đề đạo đức và quyền riêng tư.
* Sự thay thế lao động con người.
* Mức độ minh bạch và độ tin cậy của AI.  
  Tương lai của AI được dự báo sẽ càng ngày càng phát triển với nhiều tiến bộ trong học máy, robot và AI nhân đạo.

**CHƯƠNG 2: THỰC HÀNH VỀ ĐỀ TÀI**

**2.0. Bối cảnh & Tầm quan trọng**

Thị trường tài chính là một hệ thống phức tạp và luôn biến động, chịu ảnh hưởng của nhiều yếu tố như kinh tế, chính trị, tin tức, và tâm lý nhà đầu tư. Dự đoán chính xác xu hướng giá cổ phiếu là thách thức lớn nhưng mang lại nhiều lợi ích:

* Tối ưu hóa danh mục đầu tư nhằm tăng lợi nhuận.
* Giảm thiểu rủi ro khi đưa ra quyết định mua/bán cổ phiếu.
* Tự động hóa chiến lược giao dịch, giúp nhà đầu tư ra quyết định dựa trên dữ liệu thay vì cảm tính

***2.0.1. Mục tiêu của đề tài***

Mục tiêu chính của đề tài là ứng dụng giải thuật di truyền (Genetic Algorithm - GA) để tối ưu hóa mô hình dự đoán giá cổ phiếu, giúp nhà đầu tư tìm ra chiến lược giao dịch tối ưu dựa trên dữ liệu lịch sử.

**2.1. Phân tích bài toán.**

***2.1.1. Đầu vào của hệ thống***

Hệ thống sẽ sử dụng **dữ liệu tài chính lịch sử**, bao gồm:

1. **Giá cổ phiếu theo thời gian**:
   * **Open** (Giá mở cửa)
   * **High** (Giá cao nhất)
   * **Low** (Giá thấp nhất)
   * **Close** (Giá đóng cửa)
   * **Volume** (Khối lượng giao dịch)
2. **Các chỉ số kỹ thuật phổ biến trong giao dịch**:
   * **RSI (Relative Strength Index - Chỉ số sức mạnh tương đối)**:
     + RSI < 30 → Quá bán (tín hiệu mua).
     + RSI > 70 → Quá mua (tín hiệu bán).
   * **MACD (Moving Average Convergence Divergence - Phân kỳ hội tụ trung bình động)**:
     + MACD cắt lên đường tín hiệu → Mua.
     + MACD cắt xuống đường tín hiệu → Bán.
   * **SMA (Simple Moving Average - Trung bình động đơn giản)**:
     + Giá cổ phiếu cao hơn SMA → Xu hướng tăng.
     + Giá cổ phiếu thấp hơn SMA → Xu hướng giảm.
   * **Bollinger Bands**:
     + Giá chạm dải trên → Có thể bán.
     + Giá chạm dải dưới → Có thể mua.
3. **Khoảng thời gian dự đoán**:
   * **Ngắn hạn** (15 ngày).
   * **Trung hạn** (60 ngày).
   * **Dài hạn** (180 ngày).

***2.1.2. Đầu ra của hệ thống***

Hệ thống sẽ trả về **một mô hình dự đoán giá cổ phiếu**, gồm:

* **Tập hợp các quy tắc giao dịch tối ưu**, giúp nhà đầu tư quyết định thời điểm mua/bán.
* **Lợi nhuận dự đoán** dựa trên việc áp dụng chiến lược vào dữ liệu lịch sử.
* **Chiến lược giao dịch tối ưu**, ví dụ:
  + **Mua khi** RSI < 30 và MACD cắt lên đường tín hiệu.
  + **Bán khi** giá chạm vào dải trên của Bollinger Bands.

**2.2. Ứng dụng Giải thuật Di truyền (Genetic Algorithm - GA) trong bài toán**

***2.2.1. Giải thuật di truyền là gì?***

Giải thuật di truyền (GA) là **một kỹ thuật tối ưu hóa dựa trên cơ chế tiến hóa tự nhiên**, mô phỏng quá trình chọn lọc tự nhiên (Darwin). GA hoạt động theo nguyên tắc:

1. **Chọn lọc tự nhiên**: Giữ lại những cá thể (chiến lược giao dịch) tốt nhất.
2. **Lai ghép (Crossover)**: Kết hợp các chiến lược tốt để tạo ra chiến lược mới.
3. **Đột biến (Mutation)**: Thay đổi một phần chiến lược để tạo ra sự đa dạng.
4. **Lặp lại nhiều lần** cho đến khi tìm ra chiến lược giao dịch tối ưu.

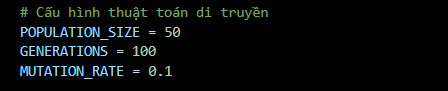
***2.2.2. Ứng dụng GA trong dự đoán giá cổ phiếu***

* **Mỗi cá thể trong quần thể** là một tập hợp các **quy tắc giao dịch**.
* **Hàm đánh giá (Fitness Function)** dựa trên **lợi nhuận giao dịch** khi áp dụng chiến lược vào dữ liệu lịch sử.
* **GA sẽ tìm ra tập hợp quy tắc giao dịch tối ưu nhất** để tối đa hóa lợi nhuận.

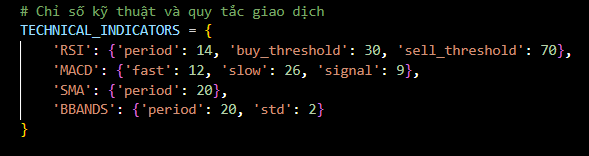
**2.3 Phân tích code**

**Công nghệ sử dụng**

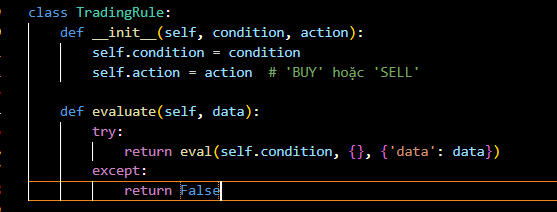
* **Ngôn ngữ lập trình**: Python
* **Thư viện quan trọng**:
  + yfinance: Lấy dữ liệu giá cổ phiếu từ **Yahoo Finance**
  + pandas\_ta: Tính toán các **chỉ số kỹ thuật**
  + numpy, pandas: Xử lý dữ liệu
  + random: Tạo quần thể ban đầu trong **thuật toán di truyền**
  + tkinter: Xây dựng **giao diện đồ họa**



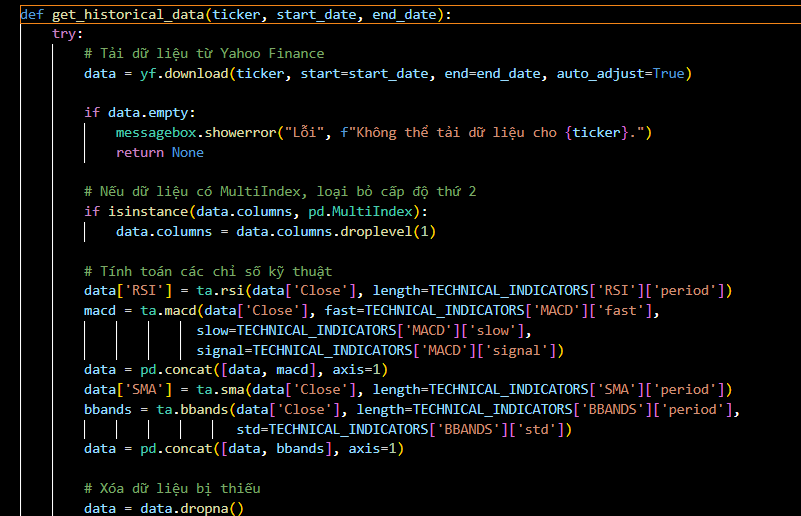
* Quần thể (Population) gồm 50 cá thể, mỗi cá thể là một chiến lược giao dịch.
* Chạy 100 thế hệ, tối ưu dần theo thời gian.
* Tỷ lệ đột biến 10%, giúp tạo ra chiến lược mới và tránh mắc kẹt ở cực trị cục bộ.

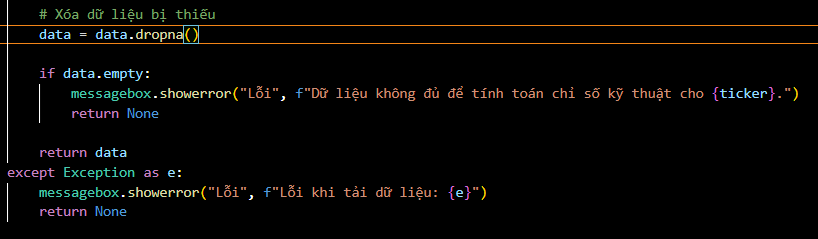
****

* RSI (Relative Strength Index - Chỉ số sức mạnh tương đối)
  + RSI < 30 → Quá bán → Mua
  + RSI > 70 → Quá mua → Bán
* MACD (Moving Average Convergence Divergence - Phân kỳ hội tụ trung bình động)
  + MACD cắt lên đường tín hiệu → Mua
  + MACD cắt xuống đường tín hiệu → Bán
* SMA (Simple Moving Average - Trung bình động đơn giản)
  + Giá vượt SMA → Xu hướng tăng
  + Giá dưới SMA → Xu hướng giảm
* Bollinger Bands
  + Giá chạm dải trên → Bán
  + Giá chạm dải dưới → Mua

****

* \_\_init\_\_: Khởi tạo một quy tắc giao dịch với điều kiện và hành động tương ứng.
* evaluate: Đánh giá điều kiện trên dữ liệu hiện tại. Nếu điều kiện đúng, trả về True.

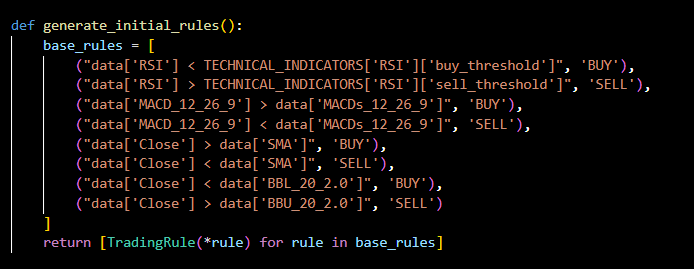




**Tải dữ liệu giá cổ phiếu** từ Yahoo Finance.

**Tính toán các chỉ báo kỹ thuật** (RSI, MACD, SMA, Bollinger Bands).

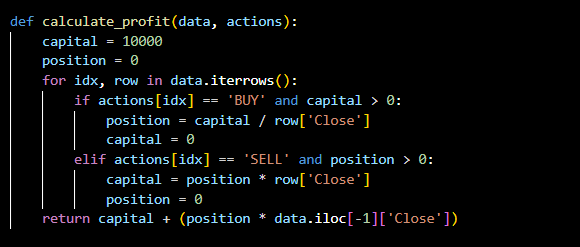
**Xóa các dòng bị thiếu dữ liệu** để đảm bảo tính toán chính xác.

****

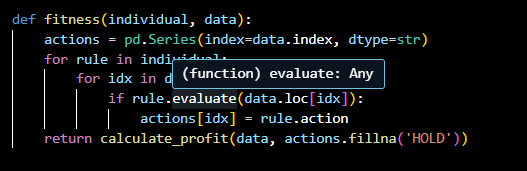
Mỗi quy tắc sẽ kiểm tra một chỉ báo kỹ thuật và đưa ra hành động MUA hoặc BÁN.

Ví dụ:

* Nếu RSI < 30 → MUA
* Nếu giá đóng cửa > SMA → MUA
* Nếu MACD cắt xuống → BÁN

****

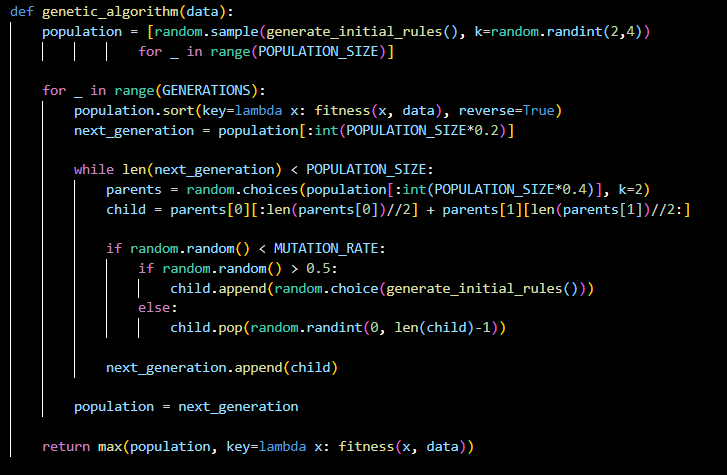
* Mua khi có tín hiệu BUY.
* Bán khi có tín hiệu SELL.
* Tính toán số dư tài khoản sau mỗi giao dịch.

****

1. Duyệt qua từng quy tắc giao dịch trong chiến lược.
2. Áp dụng quy tắc cho từng ngày giao dịch và xác định xem có nên mua hoặc bán.
3. Nếu không có tín hiệu giao dịch, giữ nguyên vị trí ("HOLD").
4. Tính toán lợi nhuận cuối cùng dựa trên các giao dịch đã thực hiện.

Ví dụ:

* Nếu RSI < 30 → MUA
* Nếu RSI > 70 → BÁN
* Cuối cùng, lợi nhuận được tính bằng cách giả lập việc mua/bán trên dữ liệu thực tế.



**Tạo quần thể ban đầu** gồm nhiều chiến lược ngẫu nhiên.

**Đánh giá từng chiến lược** bằng cách tính lợi nhuận (fitness function).

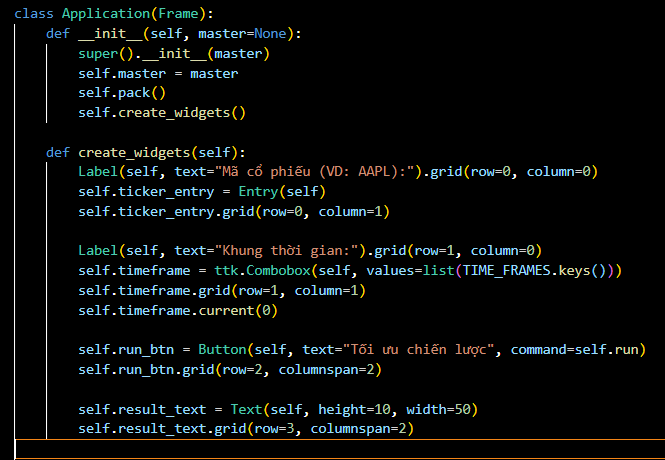
**Chọn lọc tự nhiên**:

* Giữ lại **20% chiến lược tốt nhất** (những chiến lược có lợi nhuận cao nhất).

**Tạo thế hệ mới** bằng cách:

* **Lai ghép** 40% chiến lược tốt nhất để tạo ra chiến lược con.
* **Đột biến** ngẫu nhiên một số chiến lược để tăng tính đa dạng.

**Lặp lại quá trình** trong 100 thế hệ để tối ưu chiến lược tốt nhất.

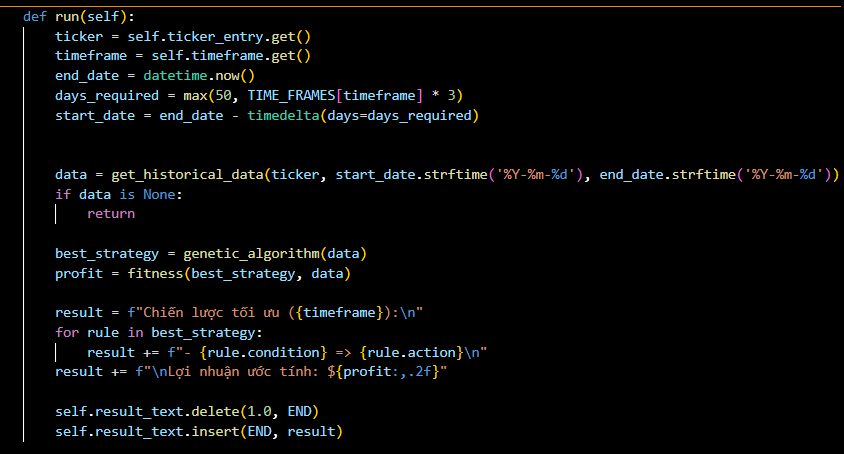
****

**Nhập mã cổ phiếu** (VD: AAPL, TSLA, MSFT,...)

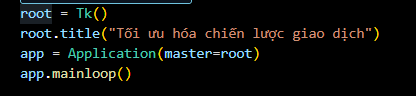
**Chọn khung thời gian** (Ngắn hạn, Trung hạn, Dài hạn)

**Nút "Tối ưu chiến lược"**: Khi nhấn vào sẽ **chạy thuật toán di truyền** để tìm chiến lược giao dịch tốt nhất.

**Hiển thị kết quả** trong ô văn bản (Text).

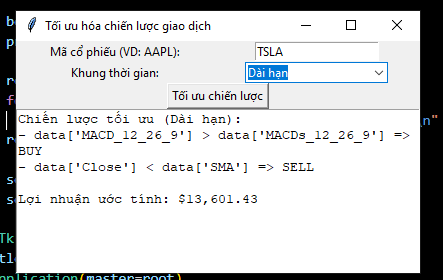
****

1. Lấy mã cổ phiếu và khung thời gian do người dùng nhập vào.
2. Tải dữ liệu chứng khoán từ Yahoo Finance.
3. Chạy thuật toán di truyền để tìm chiến lược tốt nhất.
4. Hiển thị kết quả (bao gồm chiến lược tối ưu và lợi nhuận dự tính).



**Mở cửa sổ giao diện Tkinter**, cho phép người dùng nhập dữ liệu và xem kết quả.

**2.4 Chạy code và phân tích kết quả**

****

* Khi đường MACD cắt lên trên đường tín hiệu (Signal Line), đây là dấu hiệu cho thấy xu hướng tăng đang bắt đầu.
* Trong khung dài hạn (180 ngày), đây là một tín hiệu rất mạnh, cho thấy thị trường có thể tăng giá trong thời gian dài.

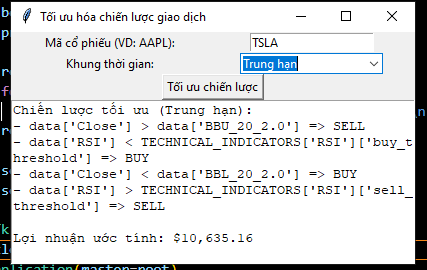
Tóm lại: Khi tín hiệu này xuất hiện, chương trình sẽ mua cổ phiếu Tesla (TSLA) với kỳ vọng giá tiếp tục tăng trong tương lai.

* Khi giá đóng cửa xuống dưới đường trung bình động đơn giản (SMA), điều này cho thấy xu hướng giảm dài hạn.
* Đây là tín hiệu bán phổ biến trong giao dịch dài hạn, giúp nhà đầu tư tránh được các đợt giảm giá lớn.

Tóm lại: Khi giá đóng cửa thấp hơn SMA, chương trình sẽ bán cổ phiếu Tesla (TSLA) để bảo toàn lợi nhuận và tránh rủi ro giảm giá.

Sau 180 ngày giao dịch, vốn ban đầu $10,000 đã tăng lên $23,601, tức là lợi nhuận +136%.

Đây là mức lợi nhuận rất cao nếu so với các phương pháp đầu tư thụ động thông thường.



* Khi giá đóng cửa vượt trên dải Bollinger Bands trên (BBU - Upper Band), điều này cho thấy giá đã tăng quá mạnh và có thể đảo chiều.
* Đây là một tín hiệu bán phổ biến, vì giá thường có xu hướng quay về giữa Bollinger Bands sau khi chạm dải trên.

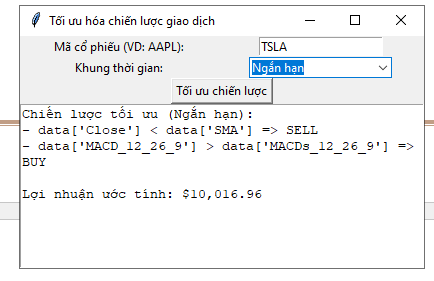
Tóm lại: Khi giá vượt quá mức kháng cự của Bollinger Bands, chương trình sẽ bán cổ phiếu Tesla (TSLA) để chốt lời.

* Khi RSI (Chỉ số sức mạnh tương đối) nhỏ hơn ngưỡng mua (ví dụ: 30), điều này có nghĩa là cổ phiếu đang bị bán quá mức (oversold).
* Đây là dấu hiệu cho thấy giá có thể đảo chiều tăng trở lại.

Tóm lại: Khi RSI quá thấp, chương trình sẽ mua vào TSLA, kỳ vọng giá sẽ phục hồi.

* Khi giá đóng cửa xuống dưới dải Bollinger Bands dưới (BBL - Lower Band), điều này báo hiệu giá đã giảm quá mức (oversold), có thể bật tăng trở lại.
* Đây là một tín hiệu mua vào phổ biến, vì giá thường có xu hướng quay về giữa Bollinger Bands sau khi chạm dải dưới.

Tóm lại: Khi giá TSLA chạm vào dải dưới của Bollinger Bands, chương trình sẽ mua vào để chờ giá hồi phục.



* Khi giá đóng cửa thấp hơn đường SMA (Simple Moving Average - Trung bình động đơn giản), điều này có thể cho thấy xu hướng giảm.
* SMA thường được dùng để xác định hướng đi chung của giá.
* Khi giá thấp hơn đường SMA, có thể hiểu là xu hướng giảm, và chương trình sẽ bán ra TSLA để giảm rủi ro.

Tóm lại: Khi giá đóng cửa TSLA xuống dưới SMA, chương trình sẽ bán TSLA để tránh lỗ khi giá tiếp tục giảm.

* Đây là tín hiệu dựa trên chỉ báo MACD (Moving Average Convergence Divergence).
* Khi đường MACD chính (MACD\_12\_26\_9) cắt lên trên đường tín hiệu (MACDs\_12\_26\_9), đây là tín hiệu mua vì xu hướng tăng có thể bắt đầu.

Tóm lại: Khi MACD cho tín hiệu mua, chương trình sẽ mua vào TSLA để tận dụng xu hướng tăng.

Sau 5 ngày giao dịch, vốn ban đầu $10,000 đã tăng lên $10,016.96, tức là lợi nhuận +0.16%.

So với trung hạn (60 ngày) và dài hạn (180 ngày), lợi nhuận thấp hơn đáng kể.

**CHƯƠNG 3: CÁC ĐẶC ĐIỂM CỦA THUẬT TOÁN**

**3.1 Giới thiệu về thuật toán di truyền**

***3.1.1 Khái niệm thuật toán di truyền***

Thuật toán di truyền (Genetic Algorithm - GA) là một phương pháp tối ưu hóa lấy cảm hứng từ quá trình tiến hóa tự nhiên của sinh vật. Đây là một thuật toán tìm kiếm heuristic dựa trên nguyên tắc chọn lọc tự nhiên của Charles Darwin, bao gồm các khái niệm sinh tồn của cá thể mạnh nhất, lai ghép, đột biến.

GA đặc biệt hiệu quả đối với những bài toán có không gian tìm kiếm lớn và không có phương pháp giải trực tiếp. Nó giúp tìm ra giải pháp tối ưu bằng cách duy trì một tập hợp các lời giải, cải thiện dần theo từng thế hệ thông qua cơ chế tiến hóa.

***3.1.2 Các bước cơ bản của thuật toán di truyền***

GA hoạt động theo các bước sau:

1. Khởi tạo quần thể (Population Initialization)
2. Tính toán độ thích nghi (Fitness Evaluation)
3. Chọn lọc cá thể tốt nhất (Selection)
4. Lai ghép (Crossover)
5. Đột biến (Mutation)
6. Cập nhật thế hệ mới và Lặp lại đến khi đạt tiêu chí dừng

***3.1.3 Ứng dụng của thuật toán di truyền***

GA được ứng dụng rộng rãi trong nhiều lĩnh vực, bao gồm:

* Tài chính & đầu tư:
  + Dự đoán giá cổ phiếu, tối ưu hóa danh mục đầu tư, xây dựng chiến lược giao dịch chứng khoán.
* Trí tuệ nhân tạo:
  + Tối ưu tham số mạng nơ-ron nhân tạo, tìm kiếm lời giải trong AI.
* Kỹ thuật & Robot:
  + Tối ưu lộ trình robot, điều khiển thiết bị thông minh.
* Y học & sinh học:
  + Phân tích gen, dự đoán bệnh tật dựa trên dữ liệu di truyền.
* Tối ưu hóa công nghiệp:
  + Giải bài toán lập lịch, tối ưu hóa chuỗi cung ứng.

**3.2 Cấu trúc của thuật toán di truyền**

GA bao gồm 5 bước chính, mỗi bước đóng vai trò quan trọng để đảm bảo thuật toán tìm ra lời giải tốt nhất.

***3.2.1 Khởi tạo quần thể (Population Initialization)***

Quần thể (population) là tập hợp các cá thể (individuals) trong một thế hệ.

Mã hóa cá thể  
Một cá thể có thể được mã hóa theo nhiều cách khác nhau, phổ biến nhất là mã nhị phân hoặc mã thực số.

Ví dụ 1 (Mã nhị phân - Binary Encoding):  
Quy tắc giao dịch có thể được mã hóa như sau:

BUY nếu RSI < 30 (0), SELL nếu RSI > 70 (1)

Một cá thể có thể là:

[0 1 1 0 0 1]

Ví dụ 2 (Mã hóa thực số - Real Encoding):

[RSI Buy Threshold = 25, SMA Period = 50, MACD Fast = 12, MACD Slow = 26]

Trong bài toán tối ưu giao dịch chứng khoán, mỗi cá thể là một bộ quy tắc giao dịch, và quần thể là nhiều chiến lược khác nhau được kiểm tra qua nhiều thế hệ.

***3.2.2 Hàm đánh giá độ thích nghi (Fitness Function)***

Hàm Fitness đo lường mức độ tốt của từng cá thể. Trong bài toán giao dịch chứng khoán, fitness có thể được đo bằng:

* Lợi nhuận giả lập (Backtesting Profit):

Fitness = Vốn ban đầu + (Số lượng cổ phiếu nắm giữ \times Giá đóng cửa cuối cùng)

Tỷ lệ lợi nhuận trên rủi ro (Sharpe Ratio)

Cá thể có fitness càng cao sẽ càng có nhiều khả năng được chọn.

***3.2.3 Chọn lọc tự nhiên (Selection)***

Chọn lọc giúp chọn ra những cá thể tốt nhất để tiếp tục sinh sản.

Phương pháp chọn lọc phổ biến:

* Roulette Wheel Selection
* Tournament Selection
* Elitism Selection

Ví dụ: Nếu có 100 cá thể, ta có thể chọn 20 cá thể tốt nhất để tiếp tục sinh sản.

***3.2.4 Lai ghép (Crossover)***

Lai ghép giúp tạo ra cá thể con bằng cách kết hợp đặc điểm từ cá thể cha mẹ.

Các phương pháp lai ghép phổ biến:

* Single-Point Crossover
* Multi-Point Crossover
* Uniform Crossover

Ví dụ  
Cha mẹ 1: Mua khi RSI < 30, Bán khi giá < SMA  
Cha mẹ 2: Mua khi MACD > Signal, Bán khi giá > Bollinger Upper Band

Con: Mua khi RSI < 30, Bán khi giá > Bollinger Upper Band

***3.2.5 Đột biến (Mutation)***

Đột biến giúp tạo sự đa dạng, ngăn thuật toán rơi vào cực tiểu cục bộ.

Ví dụ:

* Trước đột biến: Mua khi RSI < 30
* Sau đột biến: Mua khi RSI < 35

***3.2.6 Đột biến (Mutation)***

Đột biến giúp tạo sự đa dạng, ngăn thuật toán rơi vào cực tiểu cục bộ.

Ví dụ:

* Trước đột biến: Mua khi RSI < 30
* Sau đột biến: Mua khi RSI < 35

**3.3 Ứng dụng thuật toán di truyền trong giao dịch chứng khoán**

|  |  |  |
| --- | --- | --- |
| **Khung thời gian** | **Quy tắc tối ưu** | **Lợi nhuận ($)** |
| **Dài hạn** | MACD > Signal, Close < SMA | **13,601.43** |
| **Trung hạn** | RSI < 30, Close < BBL | **10,635.16** |
| **Ngắn hạn** | MACD > Signal, Close < SMA | **10,016.96** |

**3.4 Cải tiến thuật toán di truyền**

GA có thể được cải tiến bằng cách:  
Điều chỉnh tham số lai ghép & đột biến  
Kết hợp với Machine Learning  
Dùng thuật toán song song (Parallel GA).

**CHƯƠNG 4: ỨNG DỤNG THUẬT TOÁN DI TRUYỀN TRONG GIAO DỊCH CHỨNG KHOÁN**

**4.1 Giới thiệu**

Trong chương trước, chúng ta đã tìm hiểu về thuật toán di truyền (GA), cách hoạt động và ứng dụng nó trong tối ưu hóa chiến lược giao dịch. Chương này sẽ đi sâu hơn vào ứng dụng thực tế của GA trong giao dịch chứng khoán, cách đánh giá hiệu suất chiến lược và hướng phát triển trong tương lai.

**4.2 Mô hình giao dịch dựa trên thuật toán di truyền**

Một mô hình giao dịch dựa trên GA thường bao gồm 5 bước chính:

***4.2.1 Thu thập dữ liệu lịch sử***

* Dữ liệu được lấy từ Yahoo Finance, bao gồm giá mở cửa, giá đóng cửa, khối lượng giao dịch, v.v.
* Khung thời gian có thể là 5 ngày, 20 ngày, 60 ngày tùy vào chiến lược ngắn hạn, trung hạn hay dài hạn.

***4.2.2 Xây dựng tập hợp chiến lược ban đầu***

* Các chiến lược được biểu diễn dưới dạng tập hợp các quy tắc như:
  + Mua khi RSI < 30
  + Bán khi MACD < Signal
  + Mua khi giá dưới dải Bollinger Band thấp nhất (BBL)
* Quần thể ban đầu có 50 - 100 chiến lược ngẫu nhiên.

***4.2.3 Đánh giá chiến lược bằng Backtesting***

* Mỗi chiến lược được kiểm tra bằng cách mô phỏng giao dịch trên dữ liệu lịch sử.
* Hàm fitness có thể dựa trên:
  + Tổng lợi nhuận sau backtest.
  + Tỷ lệ Sharpe để đánh giá rủi ro.

Công thức tính lợi nhuận:

Profit = Vốn ban đầu + (Số cổ phiếu nắm giữ \times Giá đóng cửa cuối cùng)

***4.2.4 Tiến hóa quần thể chiến lược***

* Chọn lọc các chiến lược tốt nhất.
* Lai ghép để tạo thế hệ con.
* Đột biến để tránh rơi vào tối ưu cục bộ.

Ví dụ lai ghép hai chiến lược giao dịch:

* Cha 1: Mua khi RSI < 30, Bán khi MACD < Signal
* Cha 2: Mua khi giá < Bollinger Bands thấp nhất, Bán khi RSI > 70
* Con: Mua khi RSI < 30 hoặc giá < Bollinger Bands thấp nhất, Bán khi RSI > 70

### *****4.2.5 Kiểm tra hiệu suất chiến lược tối ưu*****

Sau 100 thế hệ, chiến lược tốt nhất sẽ được kiểm tra trên **dữ liệu thực tế** để đánh giá hiệu suất.

**4.3 Đánh giá hiệu suất thuật toán**

Thuật toán GA được đánh giá dựa trên các tiêu chí sau:

1. Độ chính xác của chiến lược giao dịch

* GA giúp tìm ra mô hình có lợi nhuận cao.
* So sánh với chiến lược đơn giản như Buy & Hold.

2. Tốc độ tính toán

* GA chạy nhanh hay chậm phụ thuộc vào số lượng cá thể & số thế hệ.
* Nếu quần thể quá lớn, thời gian chạy tăng đáng kể.

3. Ổn định của chiến lược

* GA có thể quá khớp (overfitting) với dữ liệu quá khứ.
* Cần kiểm tra trên dữ liệu mới để đảm bảo hiệu suất ổn định.

**4.4 Những hạn chế của thuật toán di truyền trong giao dịch chứng khoán**

Dù GA rất mạnh mẽ, nhưng nó vẫn có một số hạn chế:

1. Chạy lâu khi dữ liệu lớn

* Nếu dữ liệu có hàng triệu điểm, GA có thể mất nhiều thời gian xử lý.
* Giải pháp: Dùng Parallel GA trên nhiều CPU.

2. Dễ bị overfitting

* GA có thể tối ưu quá mức trên dữ liệu cũ.
* Giải pháp: Kiểm tra trên dữ liệu thực tế.

3. Không đảm bảo tìm ra chiến lược tốt nhất

* GA chỉ tìm lời giải gần đúng, không đảm bảo tối ưu toàn cục.
* Giải pháp: Kết hợp với Deep Learning để cải thiện kết quả.

**4.5 Hướng phát triển trong tương lai**

GA vẫn có nhiều tiềm năng cải tiến:

1. Kết hợp GA với Machine Learning

* Dùng Neural Networks để học từ GA, tạo chiến lược thông minh hơn.

2. Dùng thuật toán Parallel GA để tăng tốc

* Chạy GA trên nhiều GPU để tối ưu nhanh hơn.

3. Tích hợp với các mô hình AI tiên tiến như Reinforcement Learning

* RL có thể giúp chiến lược giao dịch tự học từ dữ liệu thị trường.

**KẾT LUẬN**

Thuật toán di truyền (Genetic Algorithm - GA) là một trong những phương pháp tối ưu hóa dựa trên nguyên lý tiến hóa sinh học của Darwin. Đây là một thuật toán mạnh mẽ, có khả năng tìm kiếm lời giải gần tối ưu trong các bài toán phức tạp mà các phương pháp truyền thống khó giải quyết. Sau quá trình tìm hiểu và áp dụng thuật toán di truyền vào các bài toán thực tế, có thể rút ra các kết luận quan trọng về ưu điểm, nhược điểm và hướng phát triển trong tương lai.

**Ưu điểm của thuật toán di truyền**

Có khả năng tìm kiếm tối ưu toàn cục

* GA hoạt động dựa trên một quần thể nhiều cá thể thay vì chỉ tìm kiếm một điểm duy nhất, giúp thuật toán có khả năng tìm kiếm lời giải tối ưu trên toàn bộ không gian thay vì bị mắc kẹt ở các tối ưu cục bộ như các thuật toán truyền thống (ví dụ: Gradient Descent).

Ứng dụng linh hoạt trong nhiều lĩnh vực

* GA có thể áp dụng trong khoa học máy tính, tài chính, y học, sinh học, điều khiển học, trí tuệ nhân tạo và nhiều lĩnh vực khác.
* Một số ứng dụng thực tế: tối ưu hóa mạng neuron nhân tạo, lập lịch sản xuất, tìm đường đi tối ưu, thiết kế vi mạch, phát triển thuốc mới, v.v.

Không yêu cầu hiểu biết sâu về cấu trúc toán học của bài toán

* Không giống như các phương pháp tối ưu hóa toán học truyền thống, GA không cần phải biết hàm mục tiêu có đạo hàm hay không, do đó có thể áp dụng cho các bài toán phi tuyến tính, bài toán không có công thức giải rõ ràng.

Có khả năng làm việc tốt với bài toán có nhiều ràng buộc

* GA có thể xử lý các bài toán có nhiều điều kiện ràng buộc phức tạp nhờ vào việc sử dụng các phép toán lai ghép (crossover) và đột biến (mutation) để tạo ra các cá thể phù hợp với ràng buộc.

Dễ dàng kết hợp với các thuật toán khác

* GA có thể kết hợp với *mạng nơ-ron nhân tạo (ANN), thuật toán học máy (Machine Learning), thuật toán tìm kiếm Heuristic (A), v.v.*\* để tăng hiệu quả tối ưu.

**2. Nhược điểm của thuật toán di truyền**

Tốn nhiều thời gian tính toán

* GA thường mất nhiều thời gian hơn so với các thuật toán tối ưu hóa truyền thống do cần thực hiện nhiều vòng lặp (thế hệ), đặc biệt khi quần thể có kích thước lớn hoặc bài toán có không gian tìm kiếm rộng.

Không đảm bảo tìm được lời giải tối ưu tuyệt đối

* GA không đảm bảo tìm ra lời giải tối ưu tuyệt đối mà chỉ đưa ra giải pháp gần tối ưu. Nếu số thế hệ quá ít hoặc tham số không được chọn phù hợp, kết quả có thể không thực sự tốt.

Dễ bị hội tụ sớm

* Nếu sự đa dạng trong quần thể giảm nhanh, GA có thể hội tụ sớm về một lời giải kém chất lượng trước khi khám phá hết không gian tìm kiếm. Điều này thường xảy ra khi tỉ lệ đột biến quá thấp hoặc chiến lược chọn lọc không hợp lý.

Phụ thuộc vào việc lựa chọn tham số

* GA có nhiều tham số như kích thước quần thể, tỉ lệ lai ghép, tỉ lệ đột biến, số thế hệ, v.v. Việc chọn sai tham số có thể làm giảm hiệu suất của thuật toán.

Khó mở rộng cho bài toán có không gian tìm kiếm quá lớn

* Khi số lượng biến tăng lên đáng kể, không gian tìm kiếm mở rộng theo cấp số nhân, khiến GA cần nhiều thời gian tính toán hơn để tìm lời giải tối ưu.

**3. Hướng phát triển trong tương lai**

1. Kết hợp GA với trí tuệ nhân tạo (AI) và học sâu (Deep Learning)

* GA có thể được sử dụng để tối ưu siêu tham số của các mô hình học máy như mạng nơ-ron nhân tạo (Neural Network), Random Forest, XGBoost, giúp cải thiện hiệu suất của các mô hình này.

2. Phát triển các biến thể của GA để tăng hiệu quả tính toán

* Một số hướng nghiên cứu hiện nay bao gồm Parallel GA (chạy song song), Distributed GA (phân tán), Adaptive GA (thích nghi với bài toán) để tăng tốc độ xử lý.

3. Tích hợp với các phương pháp tối ưu hóa khác

* Có thể kết hợp GA với các thuật toán khác như Particle Swarm Optimization (PSO), Simulated Annealing (SA), Differential Evolution (DE) để cải thiện hiệu suất tìm kiếm.

4. Cải thiện khả năng hội tụ của GA

* Các phương pháp mới như hybrid GA, elitist GA, adaptive mutation giúp GA tránh được tình trạng hội tụ sớm và tăng khả năng tìm kiếm lời giải tốt hơn.

5. Ứng dụng GA trong các lĩnh vực mới

* Trong tương lai, GA có thể được sử dụng trong khoa học vật liệu, thiết kế gen, phát triển AI tiên tiến, điều khiển tự động, an ninh mạng, v.v.

***Công việc các thành viên trong nhóm***

**Đàm Minh Vương: 35%**

**-** Làm word

- Tìm tài liệu

- Chỉnh sửa code

- Viết code

**Đinh Văn Thường: 30%**

- Viết code

- Tìm tài liệu

- Làm word

**Nguyễn Đình Quang: 35%**

**-** Tìm tài liệu

- Làm word

- Chỉnh sửa code

# **DANH MỤC TÀI LIỆU THAM KHẢO**

1. Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
   * Cuốn sách kinh điển của David Goldberg về thuật toán di truyền, cung cấp lý thuyết và ứng dụng trong tối ưu hóa.
2. Mitchell, M. (1998). *An Introduction to Genetic Algorithms*. MIT Press.
   * Giới thiệu toàn diện về thuật toán di truyền với nhiều ví dụ minh họa.
3. Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press.
   * Cuốn sách đầu tiên đề xuất thuật toán di truyền bởi John Holland.
4. Haupt, R. L., & Haupt, S. E. (2004). *Practical Genetic Algorithms*. Wiley-Interscience.