Materials Discovery Virtual Lab (MDVL)

Project Description

The **Materials Discovery Virtual Lab (MDVL)** is an interactive **web-based simulation** that models how scientists explore and optimize new materials using artificial intelligence principles.
Built with **Mesa (agent-based modeling)** and **Streamlit (web interface)** on **Replit**, the lab demonstrates how an **Evolutionary Strategy (ES)** algorithm can simulate *collective innovation* — where many digital "scientists" experiment, mutate, and evolve better materials over time.

Goal

To show how **agent-based modeling** and **evolutionary algorithms** can be applied to **materials discovery**, enabling participants to:

- Understand how optimization emerges from many small, local experiments.
- Visualize the effect of mutations and parameter tuning on discovery outcomes.
- Experience an **AI-driven research process** interactively through a simple web app.

Simulation Concept

Each agent represents a **virtual scientist** working on one candidate material with four measurable properties:

- **Density** — lower is better
- **Hardness** — higher is better
- **Conductivity** — higher is better
- **Cost** — lower is better

The performance of each material is evaluated with the formula:

$$\text{score} = 0.35\,(\text{hardness}_n) + 0.35\,(\text{conductivity}_n) + 0.20\,(1 - \text{density}_n) + 0.10\,(1 - \text{cost}_n)$$

Agents then apply small random **mutations** (±7%) to these properties, evaluate new scores, and gradually evolve the population toward optimal materials.

Algorithm Behind the Simulation

The app implements a simplified **Evolutionary Strategy (ES)** — a well-known optimization algorithm based on mutation, evaluation, and selection.
Each simulation cycle consists of:

1. **Evaluate** — Agents compute the performance score of their material.
2. **Mutate** — Agents slightly modify one or two features using Gaussian noise.
3. **Select & Analyze** — The system keeps track of the best-performing materials.
4. **Repeat** — Over many steps, the population collectively improves.

This process mirrors real research workflows in materials science — repeated small experiments leading to big discoveries.

System Components

COMPONENT	DESCRIPTION
DATASET	Synthetic dataset (materials_discovery_regression.csv) with ~500 candidate materials.
AGENTS	ValidatorAgent, ScientistAgents, AnalyzerAgent, ParameterAgent, VisualizerAgent.
ALGORITHM	Evolutionary Strategy (ES) implemented manually using mutation and selection.
INTERFACE	Streamlit-based web UI with sliders and live charts.
PLATFORM	Replit — fully browser-based, no installation required.

User Interface Features

- Upload or generate material datasets.
- Adjust key simulation parameters using sliders:
  - Agent count (50–500)
  - Mutation rate (0.0–0.6)
  - Steps (50–1000)
- Control simulation flow: **Run / Pause / Reset / Step once**.
- Watch results update in real time:
  - Line charts (best & mean score)
  - Diversity visualization
  - Top-10 materials table

Expected Output

- A dynamic, visual demonstration of how simple mutation and evaluation can drive optimization.
- Improved material candidates over iterations (higher scores).
- Clear understanding of **Evolutionary Strategy** and **agent-based discovery** principles.

Implementation Environment

- **Platform:** Replit (Python 3)
- **Libraries:** Mesa, Streamlit, Pandas, NumPy, Matplotlib
- **Algorithm:** Evolutionary Strategy (ES)
- **Dataset:** Synthetic dataset of candidate materials with 4 properties

Learning Outcomes

By completing this hands-on project, participants will:

- Learn how to apply **Evolutionary Algorithms** within an **Agent-Based Simulation**.
- Understand **how optimization emerges** through local agent interactions.
- Gain practical experience with **Mesa** and **Streamlit** integration on **Replit**.
- Visualize the process of **materials discovery** as an interactive AI-driven exploration.