GRDN AI Proposal

Ali El-Zein & Akshay Dhamsania
Texas A&M University

Introduction

Companion planting is a long-standing agricultural strategy that takes advantage of the inherent affinities between different plant species to boost garden productivity. This approach not only enhances soil nutrients via biological symbiosis, but it also promotes healthier crop development and serves as an organic pest control strategy. According to companion planting principles, certain plants can benefit others when planted in close proximity, either by preventing pests, increasing growth, or boosting flavor and health.

In the modern era of technology, utilizing artificial intelligence to maximize conventional techniques such as companion planting can dramatically improve their effectiveness and accessibility. GRDN.AI, a new application, exemplifies this synthesis by combining a genetic algorithm with AI to simplify and optimize companion planting. The application is intended to examine and recommend appropriate plant combinations based on a large database of plant compatibilities. This smart gardening assistant allows users to choose the sorts of plants they want to cultivate, the amount of plant beds they have, and the desired plant diversity within each bed.

The basis of GRDN.AI is its dual-backend workflow:

Mathematical Optimization Algorithm: This aspect finds the most successful plant pairings using a genetic algorithm, and this is inspired by the natural selection process. It iteratively generates the garden layout by deciding, traversing, and tweaking plant groupings based on the compatibility matrix, maximizing overall garden health and productivity.

AI-Driven Assessment and Adaptation: In addition to the mathematical model, an AI-powered system incorporates dynamic inputs and assessments. It uses machine learning to monitor garden performance data and user comments to constantly improve the plant compatibility matrix. This AI component is critical for personalizing recommendations to changing environmental conditions and individual tastes, ensuring that garden designs remain ideal over time.

Furthermore, GRDN.AI may tailor recommendations even further by taking into account local temperature data, seasonal variations, and continuing user feedback to adjust garden plans in advance. Its flexibility makes it an effective tool for both inexperienced and professional gardeners, simplifying complicated choices and reducing the trial and error that is often associated with companion planting. It helps to promote sustainable gardening by reducing the need for chemical inputs and maximizing the use of natural resources. Gardeners can create a balance between technology and nature creating an environment in which plants and technology coexist, resulting in a healthier ecology and more productive gardens.

Related Work

Retrieval-Augmented Generation (RAG) is a new concept that is built on a lot of existing research and computer science topics. In order to build a strong base in RAG, we must read some important papers by some important researchers. In this part of the essay we will describe researchers and then role in RAG . Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." This paper introduces the concept of Retrieval-Augmented

Generation and demonstrates its application across various NLP tasks. It’s a foundational paper for understanding how retrieval can augment language generation models.

Next important paper to read is by Kelvin Guu. Kevin Guu, in "REALM: Retrieval-Augmented Language Model Pre-Training." integrates retrieval into the language model pre-training process, enhancing the model's ability to utilize external knowledge effectively. Next paper is by Vladimir Karpukhin. Vladimir Karpukhin in "Dense Passage Retrieval for Open-Domain Question Answering." His work introduces a novel dense vector retrieval method that significantly improves the efficiency and accuracy of retrieving information for question answering systems.

Next significant work is by Edouard Grave and Gautier Izacard. In their paper "Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering" they explor how generative models can be combined with passage retrieval techniques to enhance open-domain question answering.

Another important paper "Generating Fact Checking Briefs." is by Angela Fan . This research focuses on using retrieval-augmented generation for creating concise fact-checking reports, demonstrating the practical application of RAG in combating misinformation.

Adam Roberts, Colin Raffel, and Noam Shazeer research paper on how much data can be passed at once is also worth a mention. Even though we have not used it extensively, its approach in problem solving is worth a consideration in our final framework of problem solving.

These papers provide a solid foundation for anyone looking to understand the scope and capabilities of RAG systems. They cover both the theoretical underpinnings and practical applications, making them ideal for beginners who wish to explore this exciting area of AI research.

Now we will talk about important concepts that are covered in our project .

- **Graph theory -** Graph theory is crucial to our project's design, helping us to smart plant through a network of plant relationships.
  - Vertices (Nodes): These are the entities in a graph. In GRDN.AI, plant species constitute the vertices of graphs.
  - Edges: Link vertices, symbolizing plant compatibilities—positive, neutral, or hostile.
  - Weight on edges: Numerical evidence of interaction strength, with values of 1, 0, or -1 for positive, neutral, or negative companion plant interactions.
  - Adjacency Matrix: A critical data matrix charting out the links and weights between all plant pairs, directing the algorithm and machine learning agent in selecting the optimal plant groups.
  - Complete Graph: A network graph arrangement in which each plant is linked to all others, guaranteeing that all relationships are considered.
- **LLMs -** We can run open-source models locally using LlamaIndex and LlamaCPP. These models include Deci's DeciLM and Meta AI's Llama2, which are used in GRDN.AI, as well as numerous other open-source models that are accessible on Hugging Face. The

fact that data stays on-premises and that all of the processing capacity of local hardware is utilized without network latency are two important advantages.

- **Learned Retrieval** - Information retrieval has seen a great deal of work on document retrieval learning, especially with pre-trained neural language models.
- **Retrieve-and-Edit approaches -** Our solution is similar to retrieve-and-edit type approaches in that it retrieves a training input-output pair identical to the input and then modifies it to produce the desired end output.

Methodology

In order to sufficiently conduct research on the effectiveness of GRDN.AI we will define the methodology used to create it. Analyzing the method in which it was built, and the results it produces, allows us to identify whether the program accomplishes its goal effectively.

The goal of GRDN AI is to use generative AI to optimize the arrangement of plants within a garden bed. The use of a genetic algorithm allows a compatibility matrix to be formed, which takes into account the nutritional needs of the plants being grown, along with other user input such as min and max # of species per plant bed.

Plants are arranged in a graph with and connected to each other with edges representing whether plantain the two crops form an overall positive, negative, or neutral relationship for each others' rate of growth. These relationships were hard coded into the program and chosen via advice given by the Farmers' Almanac.

Plant placement is also optimized by using the genetic algorithm, which takes into account population size, crossover, mutation, and seed population rates of the plants. An AI agent takes the compatibility matrix and the starting population size as parameters to produce an initial optimal arrangement of the garden bed. Other parameters such as crossover, fitness, and mutation are then used to scale up the garden bed, and produce optimal arrangements over multiple generations of plants. By minimizing the negative relationships and maximizing the positive ones, GRDN AI produces the ideal garden bed within the given parameters of the user.

References

- Fan, Angela, et al. "Generating Fact Checking Briefs." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), November 2020, pp. 7343-7360.
- Izacard, Gautier, and Edouard Grave. "Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering." Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, April 2021, pp. 874-880.
- Karpukhin, Vladimir, et al. "Dense Passage Retrieval for Open-Domain Question Answering." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), November 2020, pp. 6769-6781.
- Guu, Kelvin, et al. "REALM: Retrieval-Augmented Language Model Pre-Training." ArXiv, arXiv:2002.08909, 26 Feb. 2020.
- Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), November 2020, pp. 4912-4927.
- Kenton Lee, Ming-Wei Chang, and Kristina Toutanova. Latent retrieval for weakly supervised open domain question answering. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 6086–6096, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1612. URL https://www.aclweb.org/ anthology/P19-1612.
- https://medium.com/@dheymann314/ai-infused-optimization-in-the-wild-developing-a-companion-planting-app-357e5da29d10

- Adam Roberts, Colin Raffel, and Noam Shazeer. How much knowledge can you pack into the parameters of a language model? arXiv e-prints, 2020. URL https://arxiv.org/abs/2002.08910.
- Shiyue Zhang and Mohit Bansal. Addressing semantic drift in question generation for semisupervised question answering. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 2495–2509, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1253. URL https://www.aclweb.org/anthology/D19-1253.