Mystery Sandwich Showdown (2-4 players)

Overview

Welcome to **Mystery Sandwich Showdown**, the silliest cooking competition where expert chefs combine their wisdom to make *questionable* sandwiches!

> This game demonstrates the 'mixture-of-experts' concept in AI, where specialised inputs are weighted and integrated to form a final output in an LLM (Large Language Model - like ChatGPT). However, even if each ingredient scores highly in its own category (taste, texture, nutrition, appearance), the final combination may still be unappetising - illustrating that a system may tag an output as good even when it is not. The final discussion explores how further training and tuning can correct these mistakes.

Each player is an **expert** in one specific area:

😋 **Taste** – Decides how tasty ingredients are.
👨‍🍳 **Texture** – Evaluates crunchiness, gooeyness, or mouthfeel of the ingredients.
🍏 **Nutrition** – Decides which ingredients make the sandwich healthy (or not).
📸 **Appearance** – Judges how visually attractive each ingredient is.

And you, the facilitator will act as the **Customer** – Sets preferences and introduces chaos!

Together, the Experts must **pick ingredients, build a sandwich, and present their masterpiece**—even if it's a total disaster! At the end, they look at user preferences and discuss what a 'good' sandwich is, then use these new rules to try again and make something the Customer will actually want to eat!

Setup

Materials:

✅ **Printed Expert Cards** – Each ingredient has preset scores for **Taste, Texture, Nutrition, and Appearance.**

Team Roles:

- Get into groups of four (or pairs if necessary).
- Players choose a role from one of the four experts: **Taste, Texture, Nutrition, and Appearance.** (Choose two roles each, if in a pair.)

Gameplay Flow

1. Ingredient Selection (The Draft)

- Hand out **Expert Cards** to each participant.
- Each expert picks **TWO** ingredients from their list they think belong in the ultimate sandwich, **based on the best scores** on their card. (Not their own tastes - this represents training data of the model which will not be tailored to their personal preferences in the beginning.)

2. Sandwich Building (The Debate!)

- Players reveal their two chosen ingredients and discuss their **scores**.
- As a group, they must **pick exactly four** ingredients for their final sandwich. **No duplicate ingredients**.
- **Remind** players that they are not making a sandwich they would like, just one which gets the **highest scores in each category**. Their job is to **maximise their category score** in the finished sandwich.
- The **Customer** (facilitator) can throw in a random challenge request (see below for options)

3. The Name & Pitch (The Sell!)

- Each team **names their sandwich** (the sillier, the better).
- They **pitch it like a TV chef**, describing why it's *definitely* delicious.
- Example:
  - **"We present... The Crunchy Ocean Surprise! A daring blend of tuna, crisps, ice-cream and mustard. One bite and you'll taste the *sea... and regret!*"**
- Write each sandwich name on the board, including the total score in each category, based on its ingredients.
- Discuss each sandwich and its merits (or lack of!)
  - You (the facilitator) should start by giving your feedback about each sandwich - these can be personal judgements. "*I don't like pickles!*" This feedback will be taken into consideration later, in Step 5 - Tuning the AI.
  - If different "experts" (flavour, texture, nutrition, presentation) were making the decision, which expert would give this sandwich the highest rating? Would any expert refuse to approve it?

- Was this sandwich made by evenly weighing all expert opinions, or did one factor dominate the decision? (For example, did looks win over taste? Was weirdness more important than practicality?)
- If you got to make your sandwich again:
  - How could you make it again so the **Customer** (you, the facilitator) liked it even more?
  - Which expert would you give more weight to?
  - Would you make any special rules?

4. Judging (The Awards Ceremony)

- Which sandwich got the **highest score**? Does it seem like a nice sandwich to eat?
- Instead of picking a "best" sandwich, teams vote for fun awards (use a few of these or invent your own!):
  🏆 **Most Likely to Be Served in a Fancy Restaurant**
  😵 **Best Name**
  🤢 **Most disgusting**
  🔥 **Healthiest Bite**
  💀 **Worst Idea But Weirdly Delicious**
- Each player gets **one vote** and can't vote for their own team.

5. Tuning the AI (Aligning with user preferences)

Now that teams **know what went wrong with their sandwich**, they **get one more round** to tweak their recipe to make it more palatable to you, the customer.

- They can **set 3 new rules** (e.g., "NO peanut butter with pickles" or "Crisps are always a YES"), based on the preferences discussed previously.
- Have teams write their 3 new rules down.
- They **remake their sandwich** with the new knowledge, as per step 2.
- The **Customer** (facilitator) can throw in a random challenge request (see below for 6 options)
- This mirrors how an **AI system improves by adjusting models over time based on user feedback!**

---

6. The Re-Pitch (The Final Decision!)

Now that teams have **"fine-tuned"** their sandwich recipes based on user feedback, it's time to present their updated masterpiece—and prove that they've learned from their mistakes!

- Each team presents their new sandwich, following the same format as **Step 3: The Name & Pitch**, but with a twist: they must explain their **three new rules** and why these changes should lead to a **better sandwich model** (just like AI fine-tuning).
- The group has a discussion about how these new rules affected their process and the final output.

Presentation Steps:

1. **Announce the new sandwich name** (Has it improved? Or somehow gotten worse?)
2. **Pitch the sandwich** like a TV chef, describing why it's now the ultimate creation.
3. **Explain the three new rules** that shaped their decisions:
    - What problems did they fix?
    - Which expert gained or lost influence in the decision-making process?
    - Did their tuning make the sandwich better… or introduce unexpected side effects?

Discussion:

- Does the new sandwich **better match the customer's preferences**, or is it still flawed?
- Did any team **overcorrect**, making their sandwich **too safe or too extreme**?
- If this were an **AI model**, would their new rules make the system more useful—or just limit creativity?

Final Decision:

- After all teams have pitched, the facilitator (Customer) declares whether they are now **satisfied with the results**—or whether AI **still has a long way to go before it understands the perfect sandwich!**

Learning & Reflection: Connecting to Mixture-of-Experts AI

After the game, use these discussion questions to **help players understand how a Mixture-of-Experts (MoE) model works** in AI systems:

1️⃣ Why did some sandwiches work better than others?

💬 **Discussion Prompt:**

- Even though each expert (Taste, Texture, Nutrition, Appearance) picked high-scoring ingredients, some sandwiches were still gross. Why?
- Were there cases where one person's influence had more impact on the decision? Did that improve or weaken the final sandwich?

💡 **Mixture-of-Experts Model Connection:**

- In an MoE system, different models contribute to a final decision, but a high-quality **individual** component does not always lead to a high-quality **overall** outcome.
- If the system does not balance expert contributions correctly, the final result may be flawed—similar to how a sandwich can include good individual ingredients but still taste bad.

2️⃣ How is this like AI decision-making?

💬 **Discussion Prompt:**

- How did your group combine different perspectives to make a final decision?
- If you played again, would you prioritise certain expert opinions over others? Why or why not?

💡 **Mixture-of-Experts (MoE) Model Connection:**

- MoE architectures assign **weights** to expert models based on their relevance to the decision.
- Based on what you ask them to do, they use these weights to prioritise one expert over another to achieve the outcome you want.
- **Example**: A large language model that uses MoE (like DeepSeek) sends different parts of a user query to specialised expert networks. If a question is about coding, the model prioritises its programming expert, while a creative writing request activates a different set of experts. If the routing system fails to assign weight properly—such as favouring a factual expert for a creative writing task—the response may become boring, much like a meal that follows nutritional guidelines but tastes disgusting.

3️⃣ What rules helped make a better sandwich?

💬 **Discussion Prompt:**

- Did your team establish any rules (such as “No mixing peanut butter and tuna”)? Maybe you did, but you didn’t write them down.
- If you played again, how would you refine your approach to create a better sandwich?

💡 **Mixture-of-Experts Model Connection:**

- MoE systems improve over time through **adjustments in weighting and fine-tuning**.
- If an AI system consistently produces undesirable outputs, engineers modify training data or refine how expert models contribute.
- Your team’s decision to adjust its ingredient selection mirrors how fine-tuning in an MoE system leads to better outputs.

4️⃣ What happens if an expert is “wrong” or overconfident?

💬 **Discussion Prompt:**

- Were there any cases where one person really wanted an ingredient that ultimately made the sandwich worse?
- Did one expert’s input outweigh the others (or did the customer ask for it to do so), and if so, what was the effect?

💡 **Mixture-of-Experts Model Connection:**

- If an MoE system **over-prioritises** a particular expert model, the final decision may be biased or incorrect.
- Example: If a **safety model** in a chatbot is weighted too heavily, the chatbot may refuse to answer harmless questions.
- If a recommendation system assigns excessive weight to **popularity metrics**, it may **overlook niche but relevant content**.
- Proper weighting ensures that no single expert dominates decision-making inappropriately.

5️⃣ How could we "train" a better sandwich-making AI model?

💬 **Discussion Prompt:**

- If you could design a model that selects the best sandwich ingredients, what rules would you implement?
- How would you train it to improve decision-making over time? What feedback would you give it?

💡 **Mixture-of-Experts Model Connection:**

- MoE systems do not improve automatically—they require **training data and structured feedback**.
- If a model consistently makes poor ingredient combinations, **providing more examples of successful recipes** helps refine decision-making.
- Your team learned from experience and adjusted its selection strategy, much like how MoE models improve through iterative training and fine-tuning.

Wrap-Up: Key Takeaways

✅ Mixture-of-Experts models combine specialised systems to generate an overall decision. In this game, different experts (Taste, Texture, Nutrition, Appearance) contributed scores, just as different models in an MoE system contribute weighted inputs.

✅ High-scoring individual components do not always lead to a good final outcome. Even if an ingredient scored well in its category, the overall sandwich might not have worked. Similarly, an MoE system can produce poor results if it does not properly balance its expert models.

✅ Tuning improves outputs by refining how expert models contribute. In this game, teams adjusted their ingredient selection rules after seeing what worked. In an MoE system, adjustments to weighting and fine-tuning allow the system to generate more useful outputs.

✅ Training data and weighting determine how well an MoE system makes decisions. If the weighting is incorrect (for example, if texture is prioritised over taste), the final outcome may be undesirable. The same principle applies when training AI systems—decisions depend on how models are prioritised and trained on appropriate data.

Random Customer Challenges

Some fun and interesting ways to challenge your expert chefs with specific user requests:

1. **"VIP Customer Request!"** – A celebrity customer insists your sandwich needs a **fancy twist**! You must **add an ingredient that looks great in a picture** (high appearance score).

2. **"Oops! The Kitchen Ran Out!"** – The most delicious ingredient (high taste score) your team picked is **now unavailable**. You must replace it with **something from the lowest-scoring option** in any category.

3. **"Health Inspector Surprise!"** – The inspector is here, and they're **NOT happy**! You must **remove the least nutritious ingredient** from your sandwich and swap it for one with a high nutrition score.

4. **"The Crunch Factor!"** – A new trend is sweeping the food world: **maximum crunch!** You must swap out **your healthiest ingredient** for the best textured one available.

5. **"Double Trouble!"** – The restaurant **accidentally doubled one ingredient!** Pick one item in your sandwich and **double it**, even if it makes no sense.

6. **"The Mystery Ingredient!"** – A secret ingredient has been **mysteriously added to your sandwich**! The customer (you) should choose **a random extra ingredient** for everyone's final recipe.

7. **"My Way or the Highway!"** - Choose one of the experts and give them a higher weighting - whatever ingredients they've selected MUST be included!

Printable Expert Cards

Taste Expert

Ingredient	Score
Mayonnaise	7
Mustard	7
Hummus	9
Pesto	10
Cream cheese	5
Chicken	9
Turkey	7
Hot Sauce	4
Smoked salmon	4
Tuna	8
Cheese	8
Peanut butter	7
Jam	8
Crisps	4
Banana	9
Tomato	4
Icecream	9
Cucumber	3
Red onion	2
Pickles	6

Nutrition Expert

Ingredient	Score
Mayonnaise	7
Mustard	9
Hummus	10
Pesto	6
Cream cheese	4
Chicken	3
Turkey	4
Hot Sauce	2
Smoked salmon	7
Tuna	8
Cheese	6
Peanut butter	5
Jam	4
Crisps	2
Banana	2
Tomato	5
Icecream	1
Cucumber	5
Red onion	8
Pickles	3

Texture Expert

Ingredient	Score
Mayonnaise	3
Mustard	2
Hummus	3
Pesto	5
Cream cheese	5
Chicken	5
Turkey	7
Hot Sauce	7
Smoked salmon	2
Tuna	4
Cheese	4
Peanut butter	9
Jam	5
Crisps	10
Banana	3
Tomato	9
Icecream	6
Cucumber	5
Red onion	3
Pickles	1

Appearance Expert

Ingredient	Score
Mayonnaise	2
Mustard	1
Hummus	3
Pesto	5
Cream cheese	4
Chicken	2
Turkey	9
Hot Sauce	6
Smoked salmon	7
Tuna	3
Cheese	5
Peanut butter	1
Jam	6
Crisps	1
Banana	4
Tomato	7
Icecream	5
Cucumber	1
Red onion	10
Pickles	6

Printable Expert Cards

Taste Expert

Ingredient	Score
Mayonnaise	7
Mustard	7
Hummus	9
Pesto	10
Cream cheese	5
Chicken	9
Turkey	7
Hot Sauce	4
Smoked salmon	4
Tuna	8
Cheese	8
Peanut butter	7
Jam	8
Crisps	4
Banana	9
Tomato	4
Icecream	9
Cucumber	3
Red onion	2
Pickles	6

Nutrition Expert

Ingredient	Score
Mayonnaise	7
Mustard	9
Hummus	10
Pesto	6
Cream cheese	4
Chicken	3
Turkey	4
Hot Sauce	2
Smoked salmon	7
Tuna	8
Cheese	6
Peanut butter	5
Jam	4
Crisps	2
Banana	2
Tomato	5
Icecream	1
Cucumber	5
Red onion	8
Pickles	3

Texture Expert

Ingredient	Score
Mayonnaise	3
Mustard	2
Hummus	3
Pesto	5
Cream cheese	5
Chicken	5
Turkey	7
Hot Sauce	7
Smoked salmon	2
Tuna	4
Cheese	4
Peanut butter	9
Jam	5
Crisps	10
Banana	3
Tomato	9
Icecream	6
Cucumber	5
Red onion	3
Pickles	1

Appearance Expert

Ingredient	Score
Mayonnaise	2
Mustard	1
Hummus	3
Pesto	5
Cream cheese	4
Chicken	2
Turkey	9
Hot Sauce	6
Smoked salmon	7
Tuna	3
Cheese	5
Peanut butter	1
Jam	6
Crisps	1
Banana	4
Tomato	7
Icecream	5
Cucumber	1
Red onion	10
Pickles	6