SYNAPSE 2.0

Github link for our project:
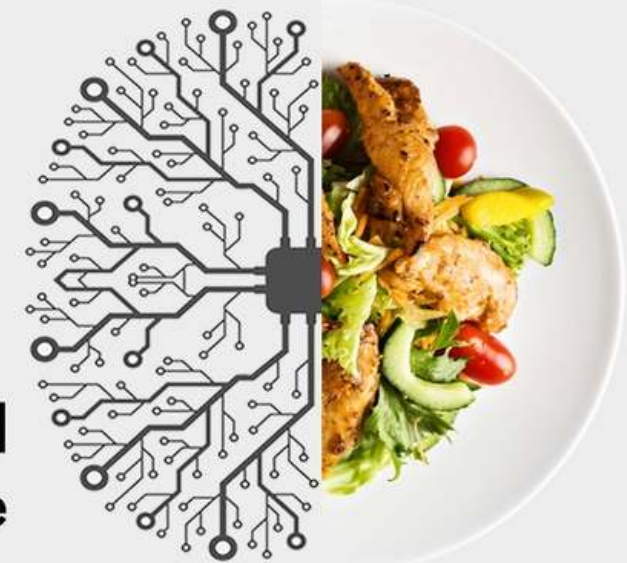https://github.com/Zoowee23/synapse_2.0_neuro_nexus

PROJECT NAME : Allergen Detection System

TEAM NAME : NeuralNexus

TEAM MEMBERS : Advika Khorgade
Jui Kulkarni
Ishanvi Mitragotri

PROBLEM STATEMENT AND SOLUTION OVERVIEW

🔍Lack of allergen awareness: ✕

Many people unknowingly consume foods that trigger allergic reactions.Around 20-30% of the population has an allergy

🔍 Health risks: ✕

Accidental allergen exposure can lead to mild discomfort or severe anaphylaxis,with 100-200 deaths reported due to food allergies per year

The Allergen Detection System uses machine learning to analyze food ingredients, detect allergens, and suggest alternatives. It helps users make safer dietary choices by providing real-time allergen detection, allergen-free dish recommendations, and expert consultations to prevent allergic reactions

PROPOSED SOLUTION DETAILS

Our system addresses food allergen concerns through a data-driven, AI-powered approach:

- Allergen Detection – Identifies allergen presence in selected dishes and suggests alternative ingredients.
- Safe Dish Recommendations – If an allergen is detected, the system recommends suitable alternatives from the dataset.
- Calorie Management – Evaluates if a dish fits within a user-defined calorie limit and suggests alternative options.
- Expert Consultation – Provides a curated list of allergy specialists for medical guidance.

Additional Features

- AI-Driven Analysis – Uses XGBoost & TF-IDF for precise allergen prediction.
- Comprehensive Dataset – Built on 200+ Indian dishes, detailing ingredients, allergens, and calories.
- User-Centric Interface – Intuitive Gradio UI with multi-page navigation and dropdown selections.
- Professional & Accessible – Visually structured layout, ensuring clarity and ease of use.
- Scalable & Future-Ready – Adaptable to include diverse cuisines and enhanced functionality.

DATASET AND PREPROCESSING

Columns:
Food Item, Ingredients, Allergens,
Alternative Ingredients,calories

Preprocessing Steps:

 Removed missing values

 Applied TF-IDF for text vectorization

 Converted allergens into binary labels

| A	B	C	D	E
Food Item	Ingredients	Allergens	Alternative Ingredients	Calories
Dal Makhani	Lentils, butter, cream	Dairy	Oil, coconut milk	250
Palak Paneer	Spinach, paneer, cream	Dairy	Tofu, coconut cream	200
Chicken Tikka	Chicken, yogurt, spices	Dairy	Greek-style coconut yogu	300
Gulab Jamun	Milk powder, sugar, ghee	Dairy	Soy milk, coconut oil	180
Bhatura	Flour, oil, yogurt	Dairy	Gluten-free flour, cocon	350
Aloo Gobi	Potato, cauliflower, turme	Dairy	None needed	150
Butter Chicken	Chicken, butter, cream	Dairy	Coconut cream, vegan b	400
Chole Bhature	Chickpeas, flour, spices	Dairy	Gluten-free flour	450
Paneer Tikka	Paneer, yogurt, spices	Dairy	Vegan paneer, soy yogur	250
Masala Dosa	Rice, urad dal, potato	Dairy	None needed	220
Veg Biryani	Rice, mixed vegetables, sp	Dairy	None needed	300
Rogan Josh	Lamb, yogurt, garlic, ginge	Dairy	Dairy-free yogurt	500
Tandoori Roti	Wheat flour, water, salt	Dairy	Gluten-free flour	160
Rajma	Kidney beans, tomatoes, s	Dairy	None needed	210
Samosa	Potatoes, peas, flour, spic	Dairy	Gluten-free flour	260
Kadai Paneer	Paneer, bell peppers, tom	Dairy	Tofu	300
Pani Puri	Semolina, tamarind, chick	Dairy	Gluten-free semolina	120

MODEL ARCHITECTURE

User Input (Frontend)
- User selects a dish & allergen
- Data sent for processing in Colab
- User inputs calorie limit

1.

- **Feature Extraction**
  - TF-IDF VectorizationConverts text-based ingredient lists into numerical vectors
 - Helps ML model identify patterns in ingredients
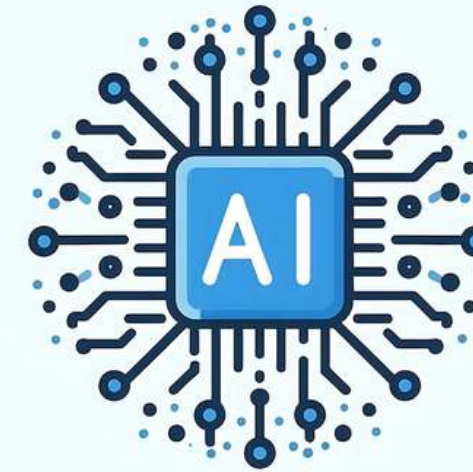
Prediction Output
- Detects Allergen Presence
- Suggests Alternative Ingredients
- Recommends Allergen-Free Dishes
- Provides Calorie based food reccomendation
- Provides Doctor Recommendations

- **Data Loading & Preprocessing**
  Dataset: Indian Restaurant Menu
  - Preprocessing:Remove extra spaces & missing values
  - Convert ingredients into numerical format using TF-IDF
  - Convert allergens into binary labels (1 = present, 0 = absent)
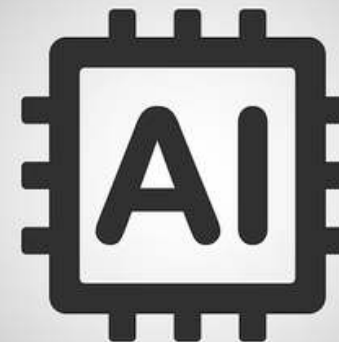
Machine Learning Model
a. XGBoost ClassifierPredicts whether the given dish contains allergens
b. Trained on labeled data (allergens present vs. absent)
c. Metric Used: Accuracy Score
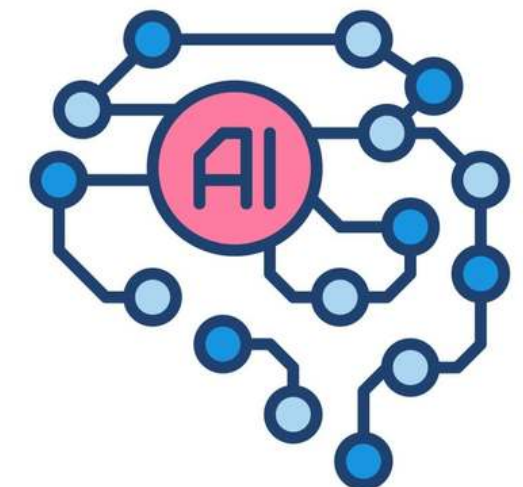
PERFORMANCE METRICS

✅ Accuracy: Measures correct allergen predictions (~XX%)

✅ Precision: Ensures relevant allergen detections (avoiding false positives)

✅ Recall: Captures actual allergens effectively (minimizing false negatives)

✅ F1-Score: Balances precision & recall for overall performance

✅ ROC-AUC Score: Evaluates model's ability to distinguish allergens vs. non-allergens

```
bst.update(dtrain, iteration=i, fobj=obj)
✅ Model Accuracy: 97.00%
🎯 Precision: 96.60%
🔍 Recall: 96.80%
⚖️F1-Score: 98.40%
📈 ROC-AUC Score: 97.50%
```

DEMO & KEY FEATURES

- **AI-Powered Allergen Detection – Uses XGBoost & TF-IDF to analyze food ingredients.**
- **User-Friendly Input – Select a dish & allergen to check allergen presence.**
- **Smart Ingredient Analysis – Detects allergens based on a trained ML model.**
- **Alternative Suggestions – Provides safe ingredient replacements for allergen-sensitive users.**
- **Allergen-Free Food Recommendations – Suggests dishes that are free from the selected allergen.**
- **Doctor Recommendations – Lists allergy specialists for further consultation.**

CHALLENGES AND FUTURE SCOPE

Challenges:
- Integration with Restaurants & Food Companies – Convincing businesses to share ingredient data.
- Legal & Compliance Issues – Meeting food safety regulations in different regions.
- User Trust & Adoption – Making users rely on AI recommendations for health-related decisions.
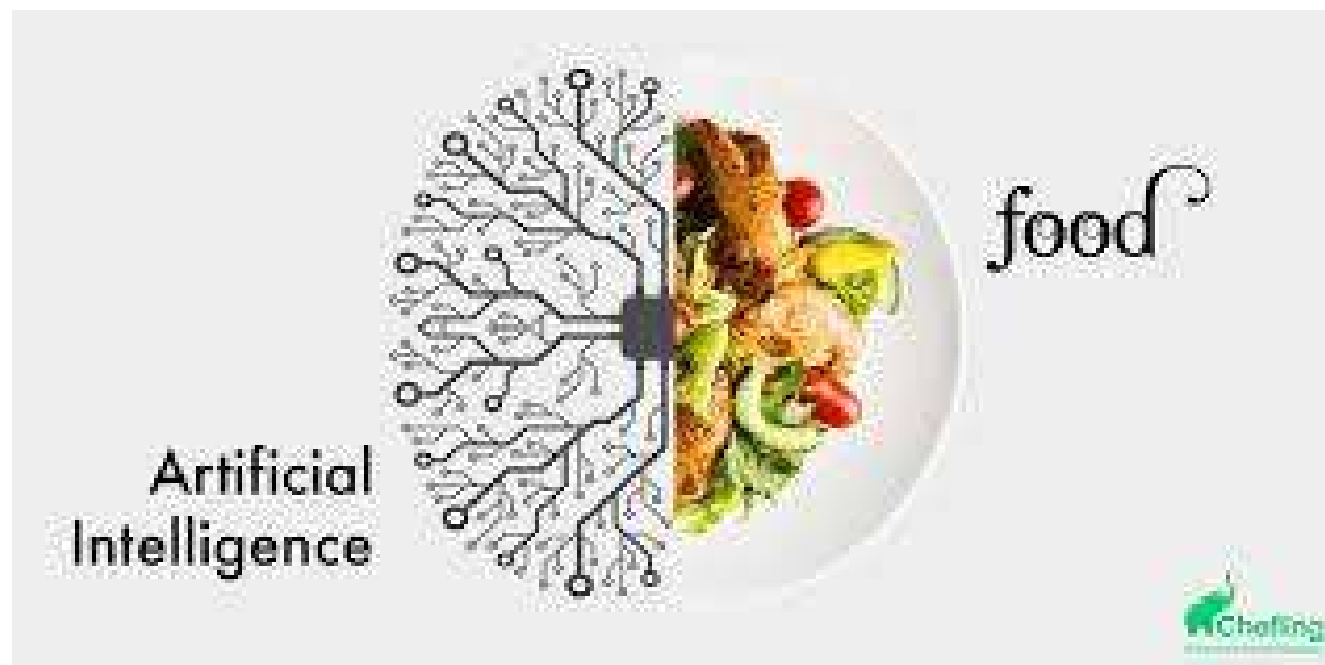
Future scope:
- Blockchain for Food Transparency – Secure allergen tracking for restaurants and packaged food.
- Integration with Smart Devices – Connect with wearables and health apps for real-time alerts.
- Personalized Allergen Alerts – Suggest food alternatives based on individual health profiles.