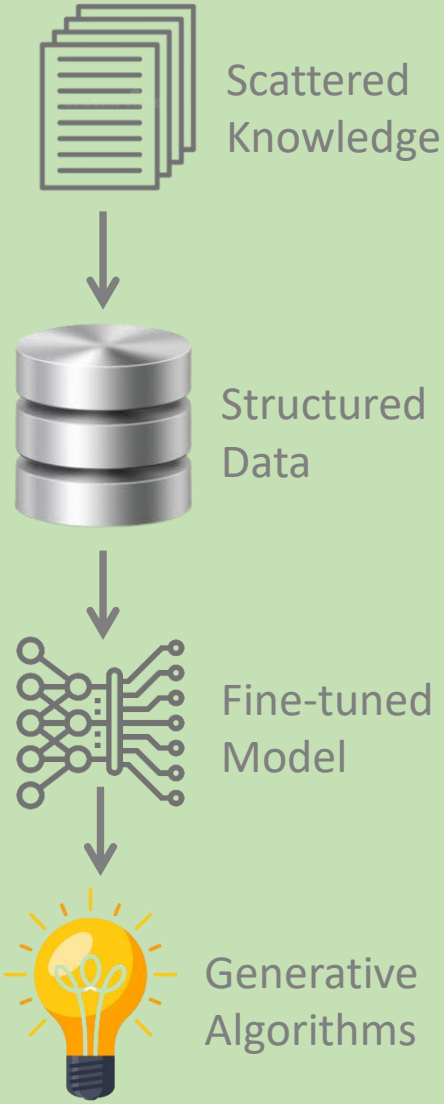
Applications

Novel Recipe Generation



Strategy for recipe generation and comparison

- Compilation of recipe data from online sources.
- Creating a structured repository (RecipeDB)
  - NER models for accurate identification of culinary entities; cubits (ingredient name, quantity, unit, form, processing, utensil)
- Implementing models for text generation
  - LSTM (Long short-term memory)
  - Character and word-level LSTMs
  - Bidirectional LSTMs
- Fine-tuning GPT2 for recipe generation

PERFORMANCE STATISTICS OF MODELS

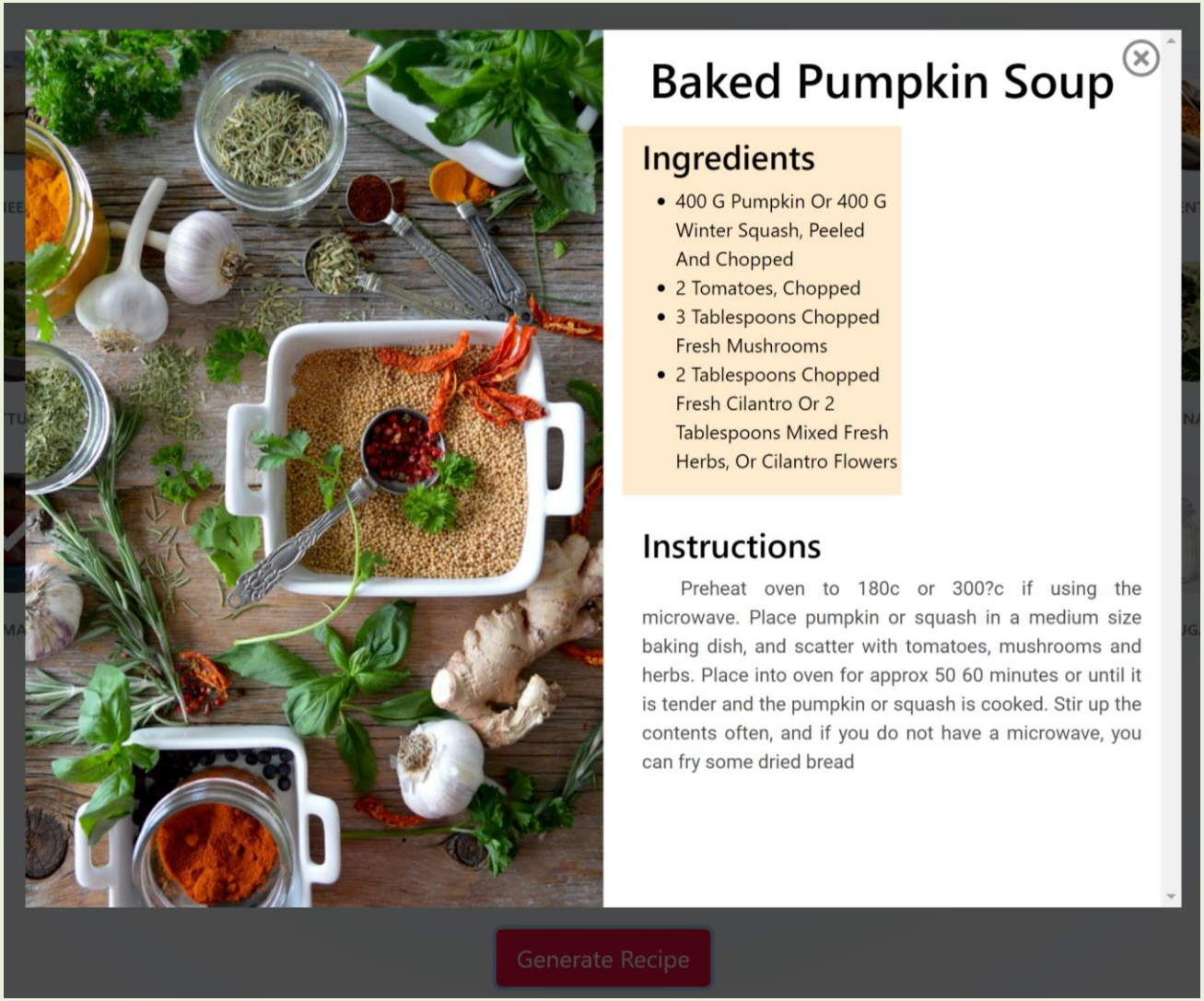
Model	BLEU Score
Char-level LSTM	0.347
Word-level LSTM	0.412
DistilGPT2	0.442
GPT-2 medium	**0.806**

Ratatouille—Can Computers Cook?

https://cosylab.iiitd.edu.in/ratatouille2

Goel et al., 'Ratatouille: A tool for Novel Recipe Generation,' IEEE 38th International Conference on Data Engineering Workshops (2022).
Y Agarwal, D Batra and G Bagler*, 'Building Hierarchically Disentangled Language Models for Text Generation with Named Entities', COLING (2020).

Ratatouille—Can Computers Cook?

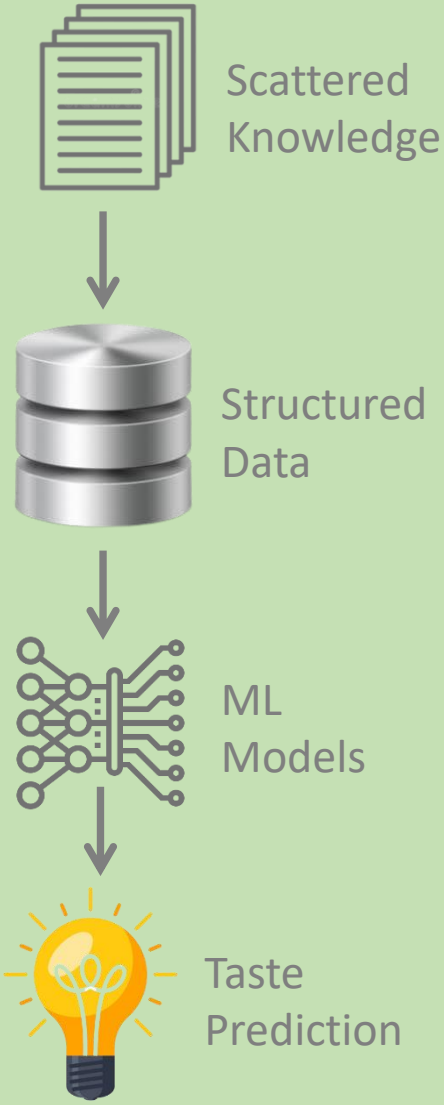


https://cosylab.iiitd.edu.in/ratatouille2

Goel et al., 'Ratatouille: A tool for Novel Recipe Generation,' IEEE 38th International Conference on Data Engineering Workshops (2022).
Y Agarwal, D Batra and G Bagler*, 'Building Hierarchically Disentangled Language Models for Text Generation with Named Entities', COLING (2020).

Applications

Taste Prediction



BitterSweet: Strategy for taste prediction

- Compilation of data of molecules and their empirically reported taste from sources.

- Creating a structured repository (BitterSweet)
  - Molecular feature extraction
  - ChemoPy, Dragon2D, Dragon2D/3D, Canvas, ECFP
  - Premise: Chemical features encode taste

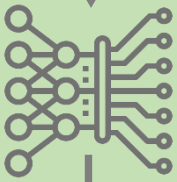
- Feature redundancy analysis

- Dimensionality Reduction
  - T-distributed stochastic neighbor embedding (t-SNE)
  - Principal Component Analysis

- Implementing classification models for taste prediction
  - Ridge Logistic Regression
  - Random Forest
  - Adaptive Boosting

Scattered Knowledge

Structured Data

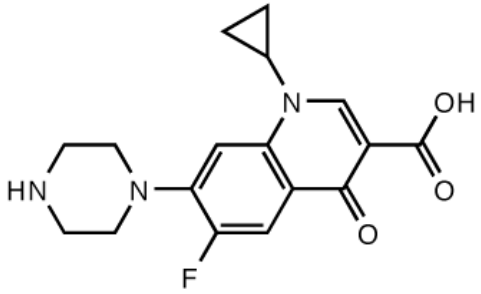


ML Models

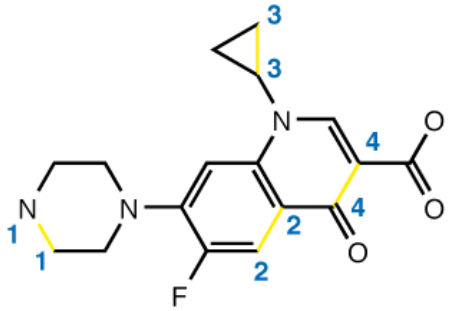
Taste Prediction

Computational representation of compounds

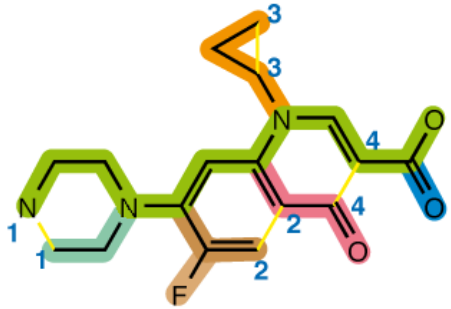
A



B



C



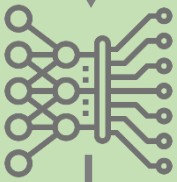
D

N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

Wikipedia

Scattered Knowledge

Structured Data



ML Models

Taste Prediction

BitterSweet: Prior research and data description

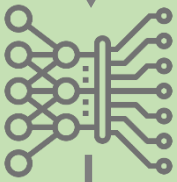
Reference	Molecular Descriptors	Taste	Source	Type	Number
BitterX[13]	Common Physicochemical Descriptors	Bitter	BitterDB[14]	Verified	539
		Non-bitter	In-house data	Verified	20
			Available Chemicals Directory	Random	519
Rojas *et al.*[16]	ECFPs and Dragon2D molecular descriptors	Sweet	Literature (Refer to supplementary materials of Rojas *et al.*[16])	Verified	336
		Bitter		Verified	81
		Tasteless		Verified	130
BitterPredict[19]	Physicochemical and ADMET descriptors	Bitter	BitterDB[14] and Rojas *et al.*[18] (bitter)	Verified	691
		Tasteless	Rojas *et al.*[18] (tasteless)	Verified	130
		Probably non-bitter	Fenaroli's Handbook of Flavor Ingredients[20]	Random	1451
		Sweet	Rojas *et al.*[18] (sweet)	Verified	336
e-Bitter[21]	ECFPs	Bitter	BitterDB[14], Rodgers *et al.*[12] and Rojas *et al.*[18] (bitter)	Verified	707
		Non-bitter	Rojas *et al.*[18] (sweet and tasteless) and BitterX[13]	Verified	149
		Sweet	SuperSweet[23], SweetenersDB[42], and Literature[18,43–45]	Verified	443
BitterSweetForest[22]	Morgan, Atom-Pair, Torsion and Morgan Feat fingerprints	Bitter	BitterDB[14]	Verified	685
		Sweet	SuperSweet[23]	Verified	517

Table 1. Overview of previous research on bitter-sweet taste prediction with details of the type of molecular descriptor used, the source of data and control(s) used, the nature and size of these data.

Tuwani et al., 'BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules,' Sci Rep 9, 7155 (2019).

Scattered Knowledge

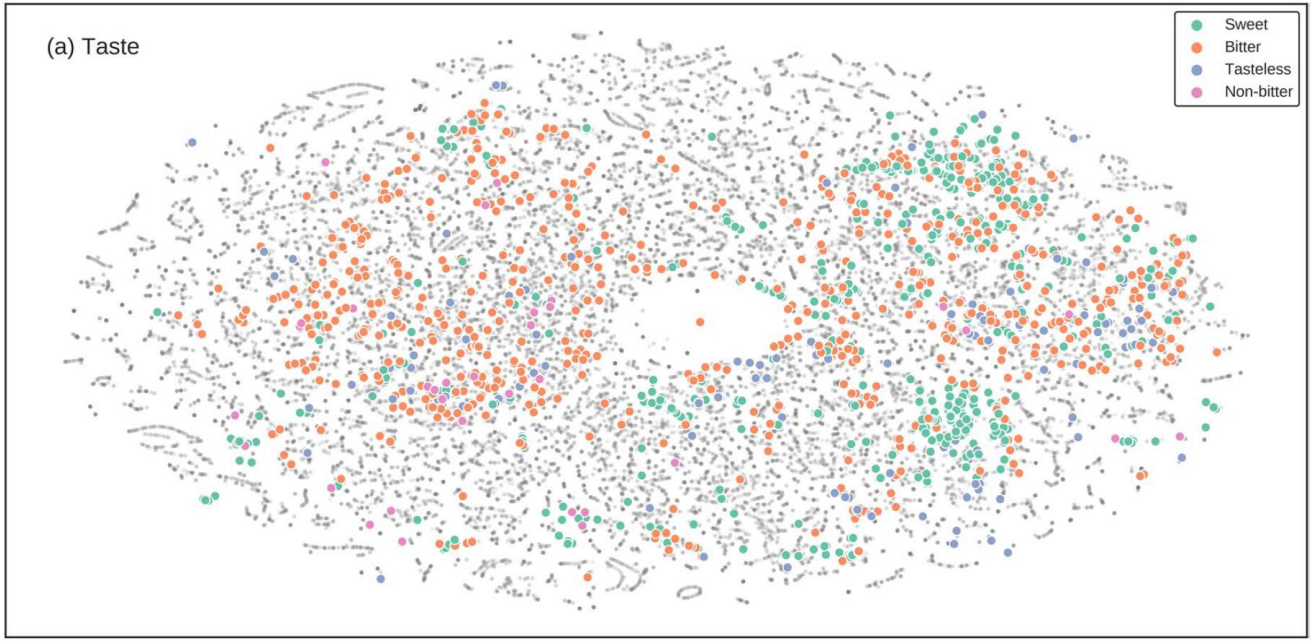
Structured Data



ML Models

Taste Prediction

BitterSweet: Structural diversity of data compiled

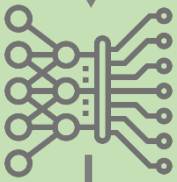


2D t-SNE scatterplot of curated and random molecules generated using physiochemical features from Canvas. Annotating taste information of molecules reveals the structural diversity of bitter, sweet and tasteless compounds as compared to random molecules.

Tuwani et al., 'BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules,' Sci Rep 9, 7155 (2019).

Scattered Knowledge

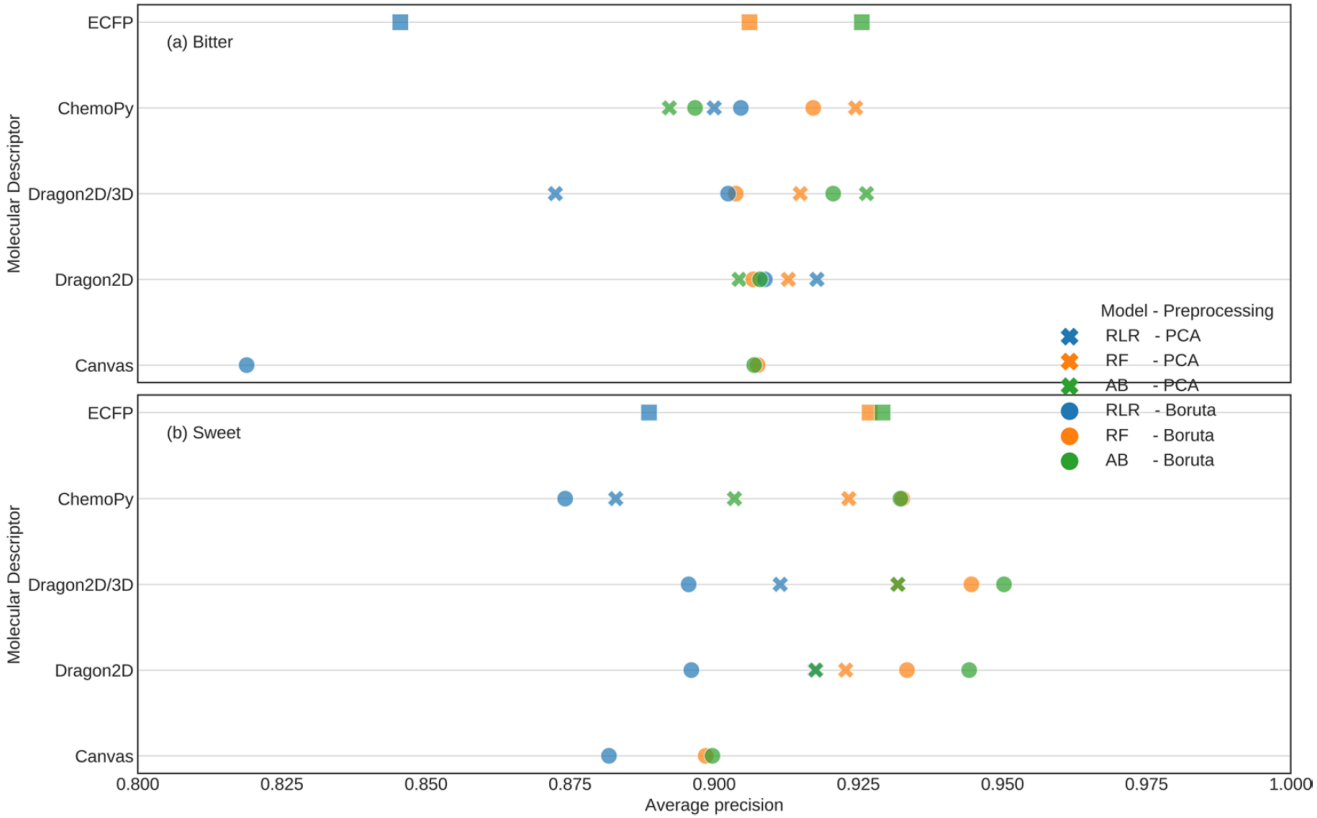
Structured Data



ML Models

Taste Prediction

BitterSweet: Performance analysis



Performance (in terms of Average Precision) of the best BitterSweet models corresponding to each molecular descriptor set for **(a) sweet/non-sweet prediction** and **(b) bitter/non-bitter prediction**.

Dragon2D/3D molecular descriptor set and Boruta feature selection were found to produce the most optimal models for sweet/non-sweet prediction.

For bitter/non-bitter prediction, PCA outperformed Boruta.

Tuwani et al., 'BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules,' Sci Rep 9, 7155 (2019).

Taste Prediction

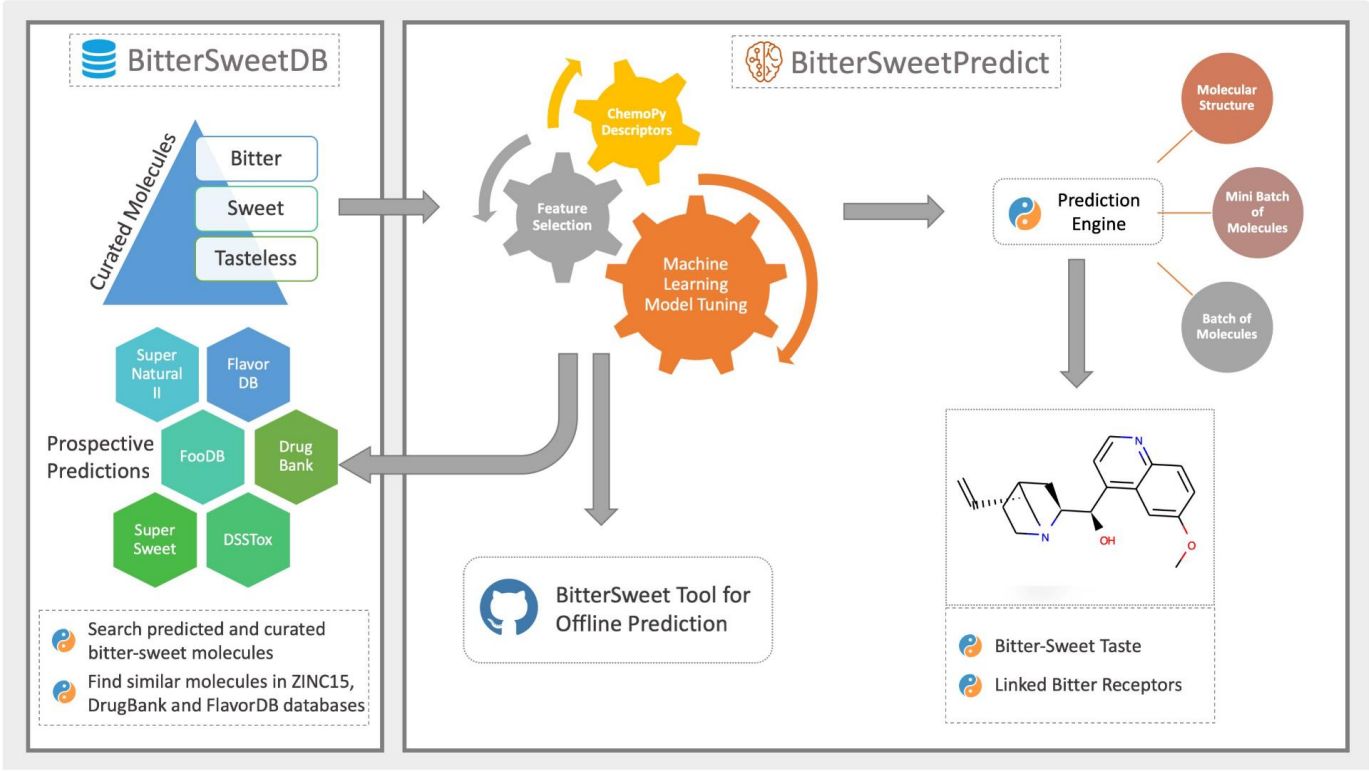
Scientific Reports (2019)

https://cosylab.iiitd.edu.in/bittersweet

BitterSweet Web Server

BitterSweet | Predict | Search | Analytics | BitterSweetPy | How to Use | FAQs | Contact Us | CoSyLab

Welcome to BitterSweet



Predict »

- Bitter and sweet tastes of small molecules
- Linked bitter receptors

Search »

- 394152 molecules with predicted BitterSweet taste
- 3086 molecules with verified BitterSweet taste

Analytics »

- BitterSweet Chemical Space
- Feature Importance(s)
- Model Performance

Server status

3 Queued jobs

0 Running jobs

1888 Completed jobs

Download BitterSweet

The BitterSweetPy Python module brings all the funtionality of BitterSweet to an offline environment, allowing you to generate predictions for an unlimited number of compounds as well as modify the machine learning pipeline to suit your needs.

Cite us

BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules (2019)

Rudraksh Tuwani, Somin Wadhwa, Ganesh Bagler

Scientific Reports 9, Article number: 7155

doi: 10.1038/s41598-019-43664-y

BitterSweet Web Server

Search Results

Search Parameters

Common Name	Quinine
Taste(s):	Bitter, Sweet, Tasteless
Source(s):	Curated
Include Predictions:	Yes

Molecules

Show 10 entries

Search:

PubChem ID	Common Name	Taste	Bitter Receptor(s)	Source(s)	Details
11920271	Quinine	Bitter	Unknown	Curated	Details
6999115	Quinine	Bitter	Unknown	Curated	Details
3034034	Quinine	Bitter	Unknown	Curated	Details
8549	Quinine	Bitter	T2R39, T2R31, T2R43...	Curated	Details
1065	Quinine	Bitter	Unknown	Curated	Details

Download

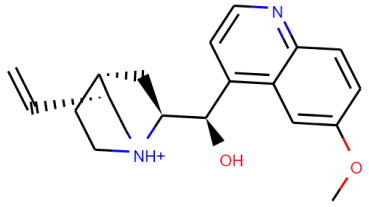
Previous 1 Next

BitterSweet Web Server

BitterSweet Profile

BitterSweet Profile

PubChem CID	11920271
Common Name	Quinine
SMILES	COC1=CC2=C(C=CN=C2C=C1)[C@H]([C@@H]3C[C@H]4CC[NH+]3C[C@@H]4C=C)O
IUPAC Name	(R)-[(2S,4R,5R)-5-ethenyl-1-azoniabicyclo[2.2.2]octan-2-yl]-(6-methoxyquinolin-4-yl)methanol
Functional Group(s)	Cation, Hydroxy compound, Alcohol, Sec. alcohol, Ether, Alkylarylether, Amine, Tert. amine, Tert. aliphat. amine, Alkene, Aromatic compound, Heterocyclic compound
☑ **Taste**	Bitter
Taste Reference:	BitterDB
Source	Curated



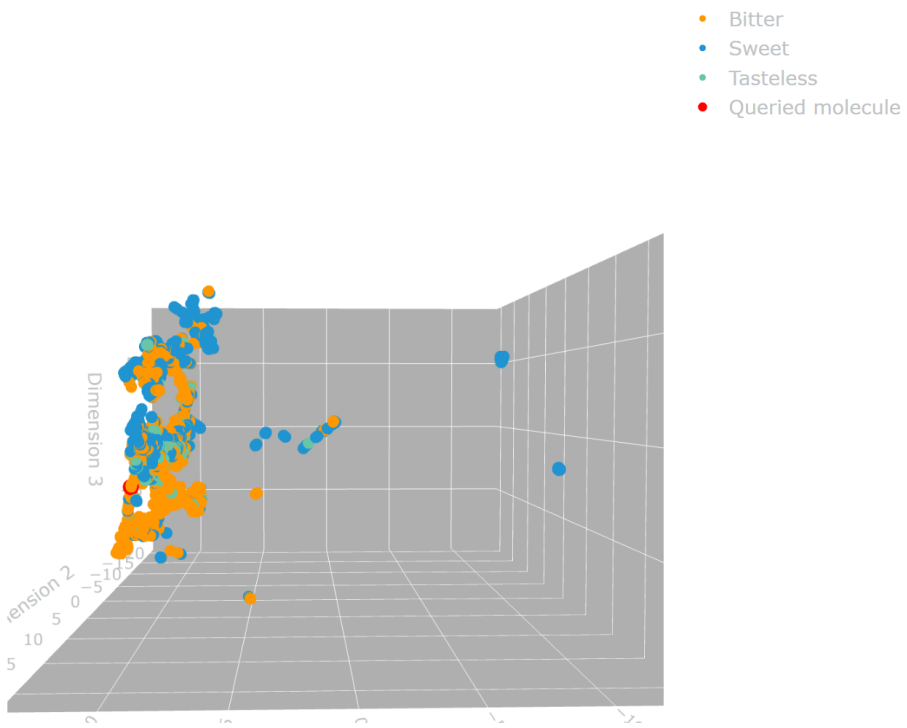
View JSmol

Search Similar in BitterSweet

Search Similar in FlavorDB

ZINC Similarity Search

UMAP View



Download

MOL File	
2D Image	
BitterSweet Profile	
Molecular Properties	

Overview Data Code Discussion Leaderboard Rules Team Host Submissions **Submit Predictions** ...

Overview Edit

Description

Evaluation

+ Add Page

The taste of an ingredient is experienced by virtue of molecules in it that interact with the taste receptors on the tongue. Sweetness is one of the major tastes to which humans have evolved attraction. Given the prevalence of type-II diabetes, it is imperative to build computational strategies for predicting whether a molecule is sweet or not, and finding ways to prediction strategies is an important challenge in computational chemistry (chemoinformatics).

Challenge

In this challenge, you'll apply your machine learning skills to predict whether a compound will be sweet or not.
You will leverage a world-class data set to build a machine learning model that challenges the current model in production. Training, validation, and testing datasets include molecular structure in SMILES format and their original taste labels. You're free to explore any technique to create the most powerful model, from creating features to using the data more organically within a model.

Acknowledgements

This dataset is part of the Complex Systems Laboratory's (IIIT-Delhi) Computational Gastronomy repository. URL: https://cosylab.iiitd.edu.in/

https://www.kaggle.com/competitions/sweet-or-not

Overview Data Code Discussion Leaderboard Rules Team Host Submissions **Submit Predictions** ...

Overview Edit

Description

Evaluation

+ Add Page

The taste of an ingredient is experienced by virtue of molecules in it that interact with the taste receptors on the tongue. Bitterness is one of the major tastes to which humans have evolved to be averse. Finding ways of predicting whether a given molecule is 'bitter or not' is an important challenge in computational chemistry (chemoinformatics).

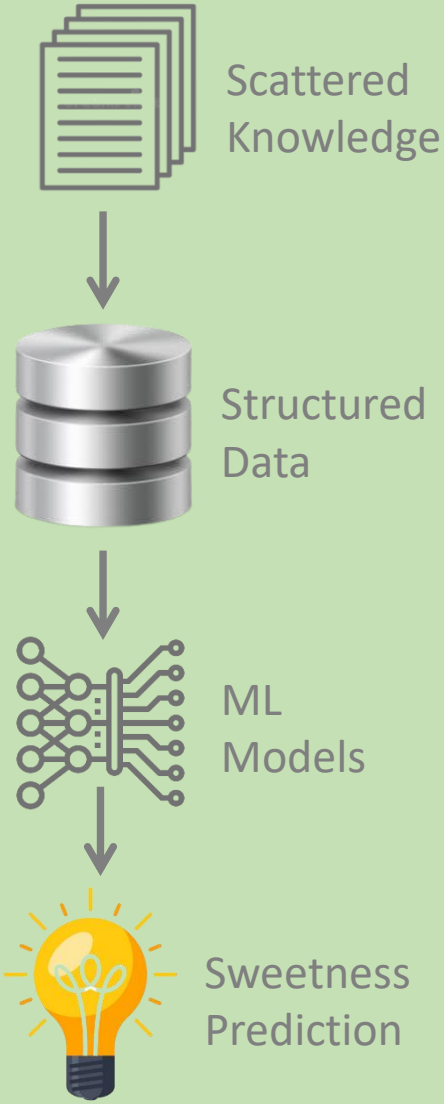
Challenge

In this challenge, you'll apply your machine learning skills to predict whether a compound will be bitter or not.
You will leverage a world-class data set to build a machine learning model that challenges the current model in production. Training, validation, and testing datasets include molecular structure in SMILES format and their original taste labels. You're free to explore any technique to create the most powerful model, from creating features to using the data more organically within a model.

Acknowledgements

This dataset is part of the Complex Systems Laboratory's (IIIT-Delhi) Computational Gastronomy repository. URL: https://cosylab.iiitd.edu.in/

https://www.kaggle.com/competitions/bitter-or-not



SweetPred: Predicting sweetness of compounds

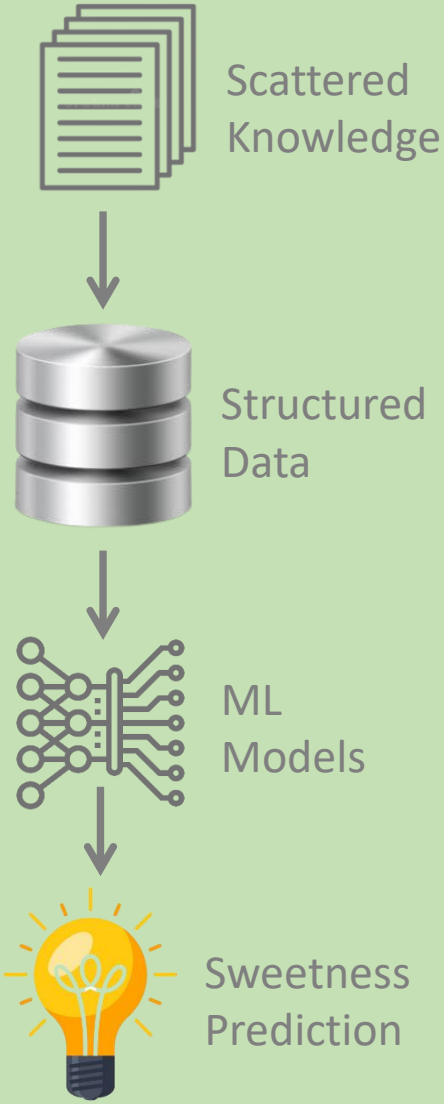
- Compilation of data of molecules and their empirically reported sweetness.

- Creating a structured repository
  - Molecular feature extraction
  - PaDel, Modred
  - Premise: Chemical features encode extent of sweetness

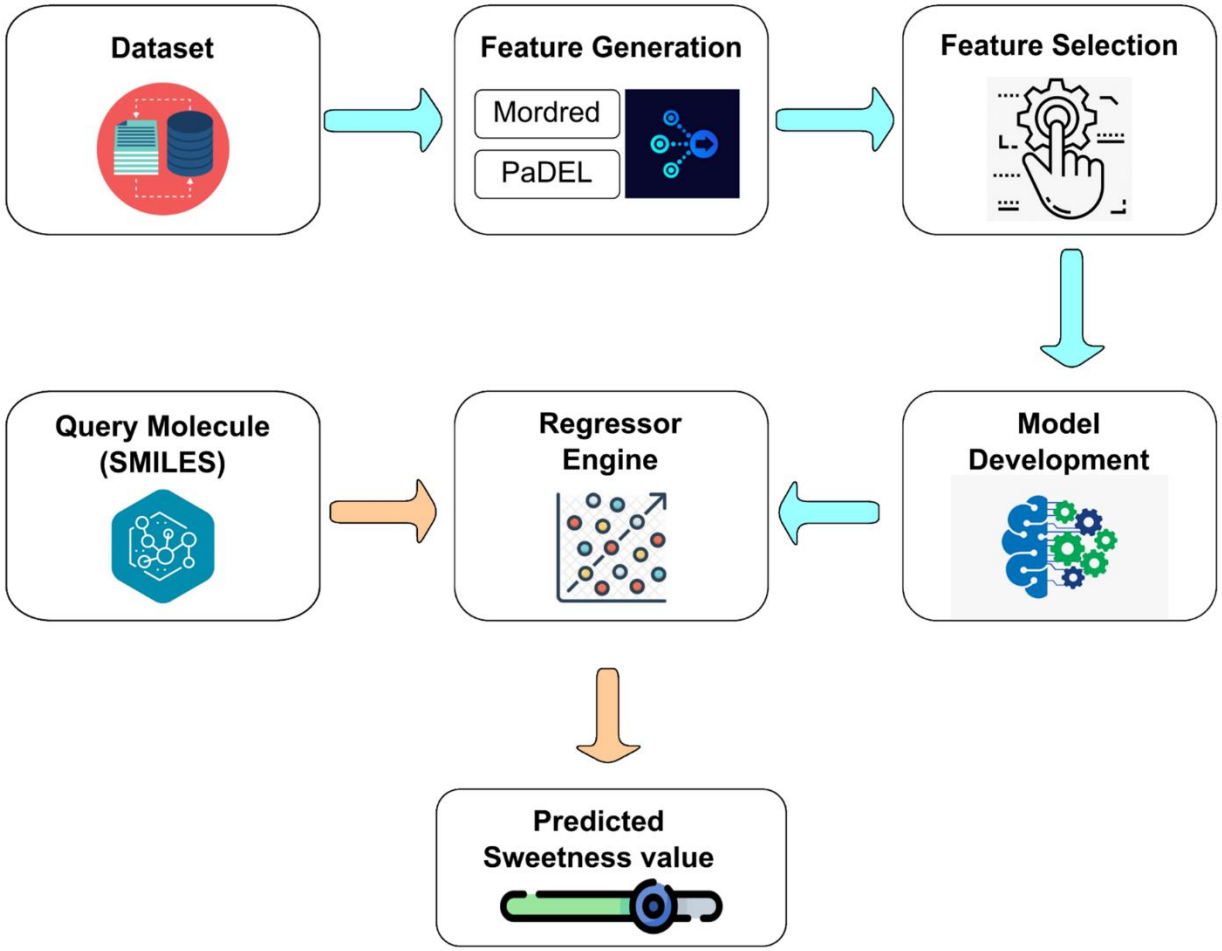
- Data pre-processing and feature selection

- Implementing regression models for sweetness prediction
  - Gradient Boost Regressor
  - Random Forest Regressor
  - Multilayer Perceptron Regressor
  - Adaboost Regressor
  - Lasso Regressor
  - Ridge Regressor
  - XGBoost Regressor

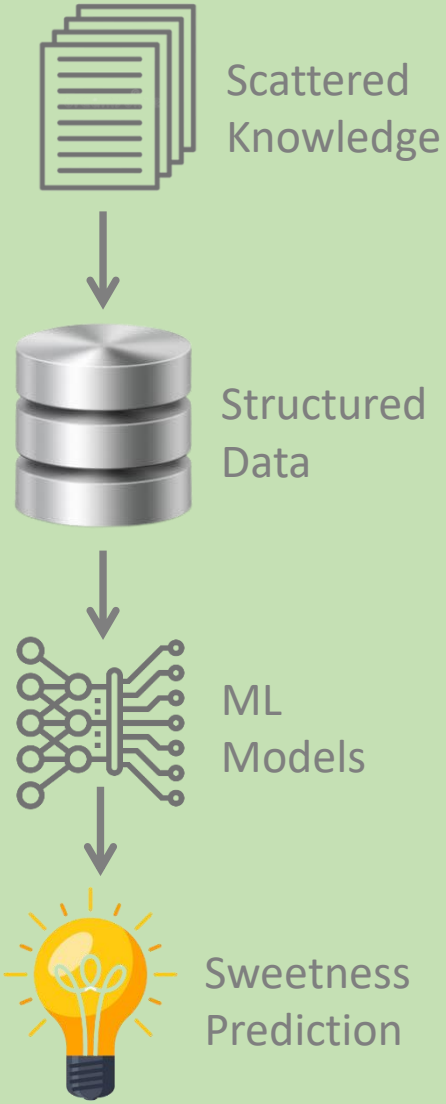
- Performance Evaluation
  - Correlation coefficient
  - Root mean square error
  - Mean absolute error



SweetPred: The schematic diagram



Goel et al., 'Machine learning models to predict sweetness of molecules,' Under review (2022).



SweetPred: Model performance

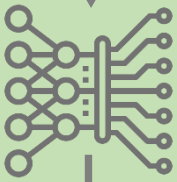
Regressor	Training Dataset			Testing Dataset		
	R	RMSE	MAE	R	RMSE	MAE
GBR	0.99	0.11	0.09	0.94	0.33	0.34
RF	0.98	0.20	0.15	0.92	0.45	0.42
AR	0.94	0.40	0.34	0.90	0.60	0.45
LR	0.96	0.30	0.20	0.87	0.54	0.45
XGB	0.90	0.68	0.64	0.83	0.80	0.75
RR	0.90	0.50	0.54	0.80	0.70	0.65
MLP	0.80	0.53	0.45	0.76	0.68	0.50

The performance statistics of machine learning models

Goel et al., 'Machine learning models to predict sweetness of molecules,' Under review (2022).

Scattered Knowledge

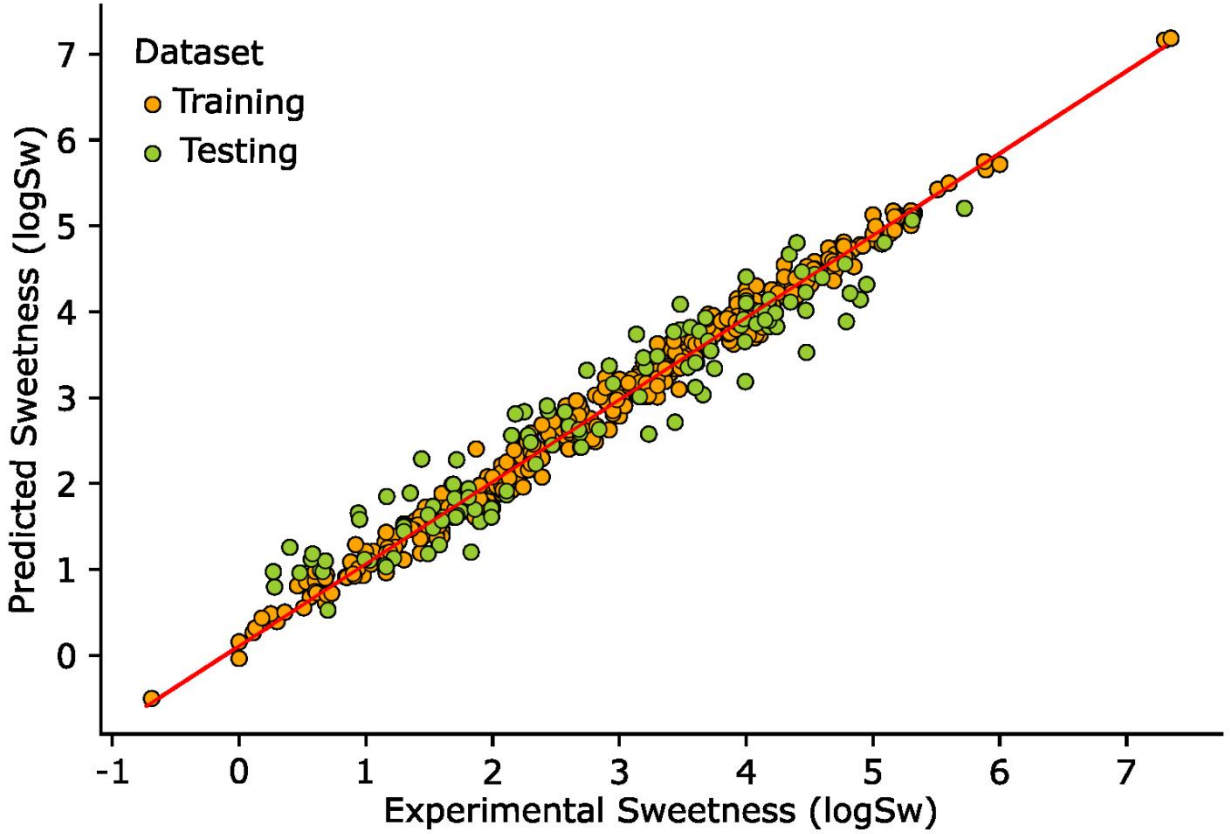
Structured Data



ML Models

Sweetness Prediction

SweetPred: Model performance



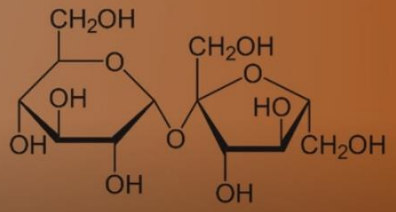
Correlation between predicted sweetness and experimental values

Goel et al., 'Machine learning models to predict sweetness of molecules,' Under review (2022).

Community Prediction Competition

How sweet is that?

Predict the sweetness value of a compound



1 teams · 3 months to go

Overview | Data | Code | Discussion | Leaderboard | Rules | Team | Host | Submissions | Submit Predictions | ...

Overview

Edit

Description

Evaluation

+ Add Page

Sweetness is an evolved human sensory mechanism that helps to identify suitable nutrients. However, with easy access to tasty and energy-dense foods, our liking for sweetness has led to an epidemic of lifestyle disorders, such as obesity, type-II diabetes, and cardiovascular issues.

Alternative sweeteners with low calorific value can make food or beverage desirable while lowering the risk of diet-linked diseases. The experimental approaches to prospect such compounds are expensive and time-consuming due to the trial-and-error involved. Staring with extensive data on sweet molecules and their experimentally reported sweetness values, machine-learning models can be of immense value in the search for suitable compounds.

Challenge

In this challenge, you will apply your machine-learning skills to predict the sweetness value of a molecule. You will leverage a world-class dataset to build a machine-learning model that gives the maximum accuracy with respect to the test set (visible and hidden). Training and testing datasets include molecular structure in SMILES format and their sweetness value. You're free to explore any technique to create the most powerful model, from creating features to using the data more organically within a model.

https://www.kaggle.com/competitions/how-sweet-is-that

Applications

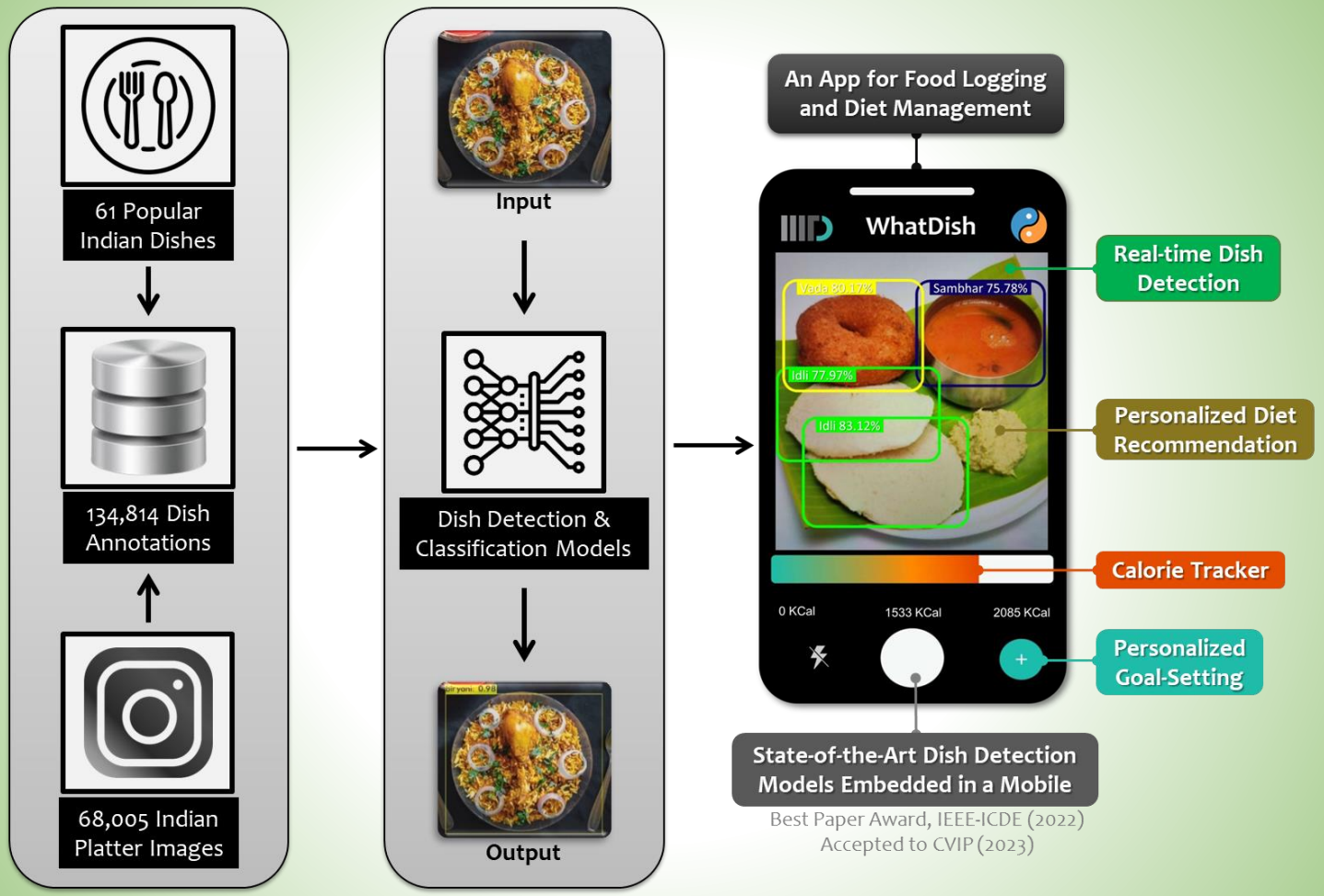
Real Time Dish Detection

Dish Detection in Indian Food Platters

- **Indian food platters** present a challenging task of accurate detection of dishes.
- We built **state-of-the-art deep-learning models** and evaluated their performance to accurately detect the **61 most popular Indian dishes**.
- **Machine learning models for accurate dish detection** have applications for **diet management** and **the mitigation of lifestyle disorders**.
- Our strategy is **extendable for dishes from across the global cuisines**.

Goel et al., “Dish Detection in Indian Food Platters: A Computational Framework for Diet Management,” **International Conference on *Computer Vision and Image Processing***, IIT Jammu, 2023 (Accepted).

Methodology



Best Paper Award, IEEE-ICDE (2022)
Accepted to CVIP (2023)

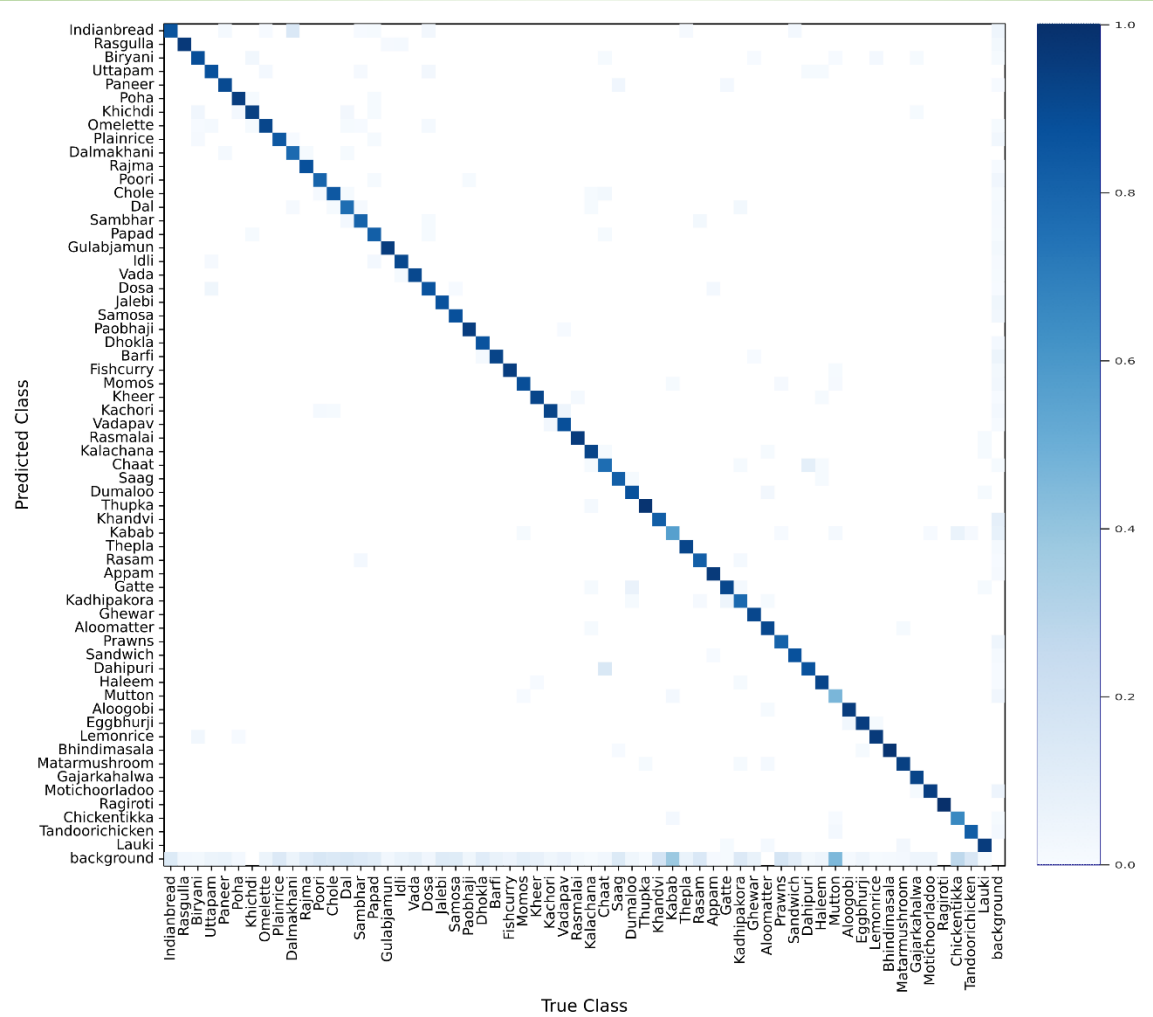
Comparison of **Multi-label Classification** Models

Classification	mAP (%) ↑	F1 score (%)	Precision (%)	Recall (%)
ResNet152	**84.51**	**88.01**	**90.56**	**85.59**
DenseNet 169	83.75	87.21	89.78	84.78
ResNet50	82.30	86.72	89.56	84.05
VGG19	77.98	83.36	88.01	79.17
DenseNet 201	77.16	82.38	86.34	78.76
VGG16	75.69	81.68	86.04	77.74
DenseNet 121	72.93	79.22	83.99	74.96
SqueezeNet1_1	57.19	68.01	81.47	58.36
SqueezeNet1_0	56.04	67.39	80.97	57.71
AlexNet	47.20	59.55	81.49	46.91

Comparison of **Object Detection** Models

Detection	mAP (%) ↑	F1 score (%)	Precision (%)	Recall (%)
YOLOv8x	**87.70**	**83.94**	**83.90**	**84.00**
YOLOv8	83.40	78.94	78.70	79.20
YOLOv7	83.50	78.92	76.60	81.40
YOLOv5	78.80	74.00	60.60	95.00
Faster RCNN	75.45	70.14	75.50	65.49
RetinaNet	71.80	72.30	71.80	72.80
DETR	52.02	60.97	52.00	73.70

Confusion Matrix



Dish Detection with YOLOv8x

Conclusions – Dish Detection in Indian Food Platters

- **Largest dataset** of annotated images of the **most popular Indian dishes** (61 classes, ~68000 images).
- **Implemented SOTA image classification & object detection models**.
- The **best model for the 'image classification' task is Resnet152** with a mAP score of 84.5%.
- **YOLOv8x emerged as the best model for 'object detection'** with mAP score of 87.7%.
- Implementing our models on IndianFood10 [Goel et al.], YOLOv8x emerged as best model with mAP of 95.55% in contrast to 91.80% by YOLOv4 (best model reported earlier).

Goel et al., "Object Detection in Indian Food Platters using Transfer Learning with YOLOv4,"
IEEE 38th International Conference on Data Engineering Workshops, Malaysia (2022).

IEEE ICDE (2020) — Best Paper Award | Computer Vision and Image Processing (2023)

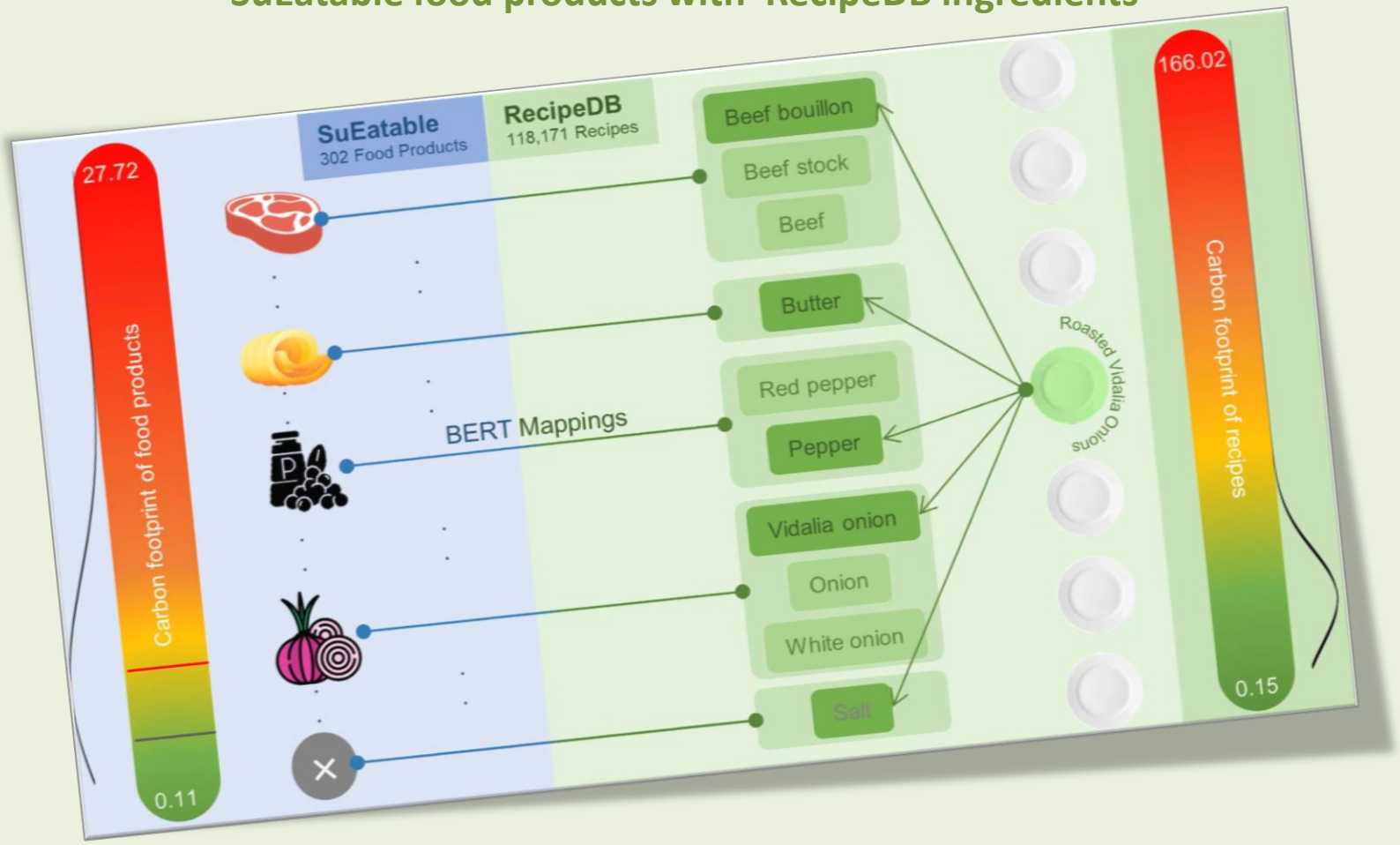
Applications

Carbon Footprint of Recipes


Food Apocalypse?

Feeding 10 Billion

Integration of carbon footprint of SuEatable food products with RecipeDB ingredients



https://cosylab.iiitd.edu.in/SustainableFoodDB/

https://cosylab.iiitd.edu.in/**SustainableFoodDB**

Food: Industry, Ethics, Policy, & Entrepreneurship

- Making products better,
- Being consumer specific
- Selling more

- Sweetness and diabetes
- Children and enticing products
- Taste and nutrition

- FSSAI
- FDA

- Data-driven food innovations