Projects

Climate change in the news

Climate change in the news

The U.S. Has Billions for Wind and Solar Projects. Good Luck Plugging Them In.

An explosion in proposed clean energy ventures has overwhelmed the system for connecting new power sources to homes and businesses.

Plans to install 3,000 acres of solar panels in Kentucky and Virginia are delayed for years. Wind farms in Minnesota and North Dakota have been abruptly canceled. And programs to encourage Massachusetts and Maine residents to adopt solar power are faltering.

The energy transition poised for takeoff in the United States amid record investment in wind, solar and other low-carbon technologies is facing a serious obstacle: The volume of projects has overwhelmed the nation's antiquated systems to connect new sources of electricity to homes and businesses.

So many projects are trying to squeeze through the approval process that delays can drag on for years, leaving some developers to throw up their hands and walk away. PJM Interconnection, which operates the nation's largest regional grid, stretching from Illinois to New Jersey, has been so inundated by connection requests that last year it <u>announced a freeze on new</u> <u>applications</u> until 2026, so that it can work through a backlog of thousands of proposals, mostly for renewable energy.

Sign up for the Climate Forward newsletter, for Times subscribers only. Your must-read guide to the climate crisis. <u>Get it in your inbox.</u>

It now takes roughly four years, on average, for developers to get approval, double the time it took a decade ago.

And when companies finally get their projects reviewed, they often face another hurdle: the local grid is at capacity, and they are required to spend much more than they planned for new transmission lines and other upgrades.

"The biggest challenge is just the sheer volume of projects," said Ken Seiler, who leads system planning at PJM Interconnection. "There are only so many power engineers out there who can do the sophisticated studies we need to do to ensure the system stays reliable, and everyone else is trying to hire them, too."

Projects Overview

The class project is your chance to go deeper into a question, dataset, and/or method you are interested in.

You will be evaluated on your ability to pose a good machine learning question in the domain of climate change (mitigation or adaptation), selection and execution of relevant methods, ability to work on a team, and your presentation skills both written and oral.

Project Timeline

March 10th - Submit project teams (5pm)

March 23rd - Written project plan due (Midnight)

March 28th - Project plan presentations in class

April 27th - Project work day in-class

May 2nd and 4th - Project presentations in class

May 5th - Project report due (Midnight)

Project teams

Everyone must work in a team (no individual projects)

Teams can be 2 or 3 people. Teams with 3 people will be expected to achieve more than teams with 2 people.

A single grade will be given to the team.

You can use the Brightspace discussion forum to organize your teams. Respond to my team member post with your team members by Friday at 5pm

Short project plan

~500 words. Provide team members and team name. Describe the problem you are tackling and how it will help mitigate or adapt to climate change and why you chose this problem. Explain the data you intend to use, the methods, and how you will evaluate your results. List anticipated challenges (e.g., data cleaning issues, learning to use new methods/coding packages, etc).

Due March 23. I will provide feedback on your plan before your presentation.

Project plan presentation

~3 minutes, 3 slides. (time yourself beforehand to make sure you have the right length!)

Explain the problem, dataset, and planned methods. Include why you've chosen what you have.

Not all team members need to speak in the project plan presentation.

March 28th in class.

Project report

Abstract - a one paragraph summary of the full report

Background - what problem are you addressing, how does it relate to climate change, how has it been approached in the past, how are you approaching it and why

Methods - what data are using, how did you process it, what methods are using, how did you decide hyperparameters, how are you evaluating your model

Results - what did your methods show, what worked and what didn't work, includes pointers to figures

Discussion - what do your results mean, what would you do in the future if you were to continue working on this project, how do your results relate to other studies

References - bibliography of sources you used and cited

~3 pages min for first five sections, plus extra pages for references and figures. Plus pre-run python notebook (should replicate your figures from data). Due May 5th by midnight

Project Presentations

~10 minutes. Summary of your project report. Should include all the same sections.

All team members must speak in the project presentation.

There will be time for questions from students and me

May 2nd and 4th in class (order will be randomly assigned)

Project timeline

Teams cannot be changed past the deadline

General project details cannot be changed after final feedback from me (after project plan presentations)

Smaller details can change as you work through your problem (message me if you don't know if the change you're considering is large or small)

Projects

Projects need to go beyond what we have done in homework assignments.

Possible approaches:

- replicate/extend an existing paper
- contribute to an established competition/benchmark
- Identify an underexplored problem where ML could help and try to solve it

Projects

Rough guidelines for creating a substantial project:

- Compare multiple methods and/or hyperparameters
- Use at least one method that isn't in scikit-learn
- Explore feature importance
- Create your own dataset from multiple sources
- Use multiple evaluation methods
- Include advanced visualizations and/or a dashboard
- etc

Here are locations to look for relevant data. Not all datasets from these sources will be climate change-related.

Note: these are real datasets. They may have missing data, poor documentation, or other challenges. You will probably have to teach yourselves some new skills to get what you want from them. These are all things data scientists have to face! Look into a dataset before committing to using it.

https://opendata.cityofnewyork.us/

Has an "environmental" section. Contains data on trees, air quality, energy use, sea level, etc.

	NY		PpenData Home Data About - Learn - Alerts Contact Us Blo	g Q Sign In
	Q, climat	e		
Categories		~	60 Results	Sort by Most Accessed 🗸 🗸
Business			Sea Level Rise Maps (2050s 500-year Floodplain) Environment	& Map
City Governme	nt		This is the 500-Year Floodplain for the 2050s based on FEMA's Preliminary Work Map data and the New York Panel on Climate Change's 90th Percentile Projects for Sea-Level Rise (31 Inches). Please see the Disclaimer PDF for more More	Updated September 16, 2021 Views 33,366
Environment			Tags sea level rise maps (2050s 500-year floodplain), sea, flood, ocean, floodplain	
Health			Sea Level Rise Maps (2020s 100-year Floodplain) Environment	& Map
Show	r AIL	~	This is the 100-Year Floodplain for the 2020s based on FEMA's Preliminary Work Map data and the New York Panel on Climate Change's 90th Percentile Projects for Sea-Level Rise (11 inches). Please see the Disclaimer PDF for more More	Updated September 16, 2021 Views 21768
Data Lens pag	es		Tags sea, flood, ocean, oltps, sea level rise maps (2020s 100-year floodplain)	
Datasets			Natural Gas Consumption by ZIP Code - 2010 Emironment	🟟 Dataset
External Datas	ets		2010 Natural Gas consumption in therms and GJ, by ZIP code, building type, and utility company.	Updated
Files and Docu	iments		Tags gas, planning, power, energy, utilities, and 1 more AP	May 9, 2022 Docs Views 10.067

Energy information administration

https://www.eia.gov/totalenergy/data/an nual/index.php

SEE ALL REPORTS Annual Energy Review Superseded see MER for key annual tables	- Annual Er Year
EIA has expanded the <i>Monthly Energy Review</i> (MER) to include annual data as far back as 1949 for those data tables that are found in both the <i>Annual Energy Review</i> (AER) and the MER. In the list of tables below, grayed-out table numbers now go to MER tables that contain data series for 1949 forward. New INTERACTIVE tables and graphs have also been added.	∍r vf
DATA CATEGORIES + EXP.	AND ALL
+ Energy overview	
+ Energy consumption by sector	
+ Financial indicators	
+ Energy resources	-11
+ Petroleum and other liquids	
+ Natural gas	
+ Coal	
+ Electricity	
+ Nuclear energy	
+ Renewable energy	
+ Environment	

+ Appendices (heat content, conversion factors, and more)

Bottom-up emissions estimations from NEI

https://www.epa.gov/air-emissions-in ventories/national-emissions-invento rv-nei

Air Emissions Inventories

	Air Emissions Inventory Home
l	Data and Documentation
	2020 NEI Data
	2017 NEI Data
	2014 NEI Data
	2011 NEI Data
	2008 NEI Data

National Emissions Inventory (NEI)

The National Emissions Inventory (NEI) is a comprehensive and detailed estimate of air emissions of criteria pollutants, criteria precursors, and hazardous air pollutants from air emissions sources. The NEI is released every three years based primarily upon data provided by State, Local, and Tribal air agencies for sources in their jurisdictions and supplemented by data developed by the US EPA. The NEI is built using the <u>Emissions Inventory System</u> (EIS) first to collect the data from State, Local, and Tribal air agencies and then to blend that data with other data sources.

CONTACT US

ClimateTRACE emissions data

https://climatetrace.org/downloads

DOWNLOADS

Climate TRACE emissions data is free and publicly available for download below. Each download package includes annual country-level emissions by sector and by greenhouse gas from 2015-2021, the applicable inventory of facility-level emissions, and facility-level ownership data where available.

VIEW DOWNLOADS BY SECTOR COUNTRY					
SECT	OR	DOWNLOAD	METHODOLOGY		
ō	Agriculture	CSV	Download (18 MB)		
Ē	Buildings	CSV	Download (0.8 MB)		
أنه	Fluorinated gases	CSV	Download (0.8 MB)		
tr.	Fossil fuel operations	CSV	Download (4 MB)		
(Q)	Forestry and Land Use	CSV	Download (0,2 MB)		
	Manufacturing	CSV	Download (15 MB)		
2	Mineral Extraction	CSV	Download (2 MB)		
ÿ	Power	CSV	Download (0.7 MB)		
6	Transportation	CSV	Download (12 MB)		
<u>M</u>	Waste	CSV	Download (7 MB)		

Carbon dioxide data from remote sensing

https://disc.gsfc.nasa.gov/d atasets/OCO2GriddedXCO 2_3/summary?keywords=0

co-2

nouncement: NASA's Terra, Aqua, and Aura Data Continuity Workshop RFI∣Aqua, Aura, Terra Near Real Time processing

Back to search results

Making Earth System Data Records for Use in Research Environments

OCO-2 Gridded bias-corrected XCO2 and other select fields aggregated as Level 4 daily files (OCO2GriddedXCO2)



Gridded carbon dioxide mole fraction (XCO2) and other select variables created by applying local kriging (also known as optimal interpolation) to daily aggregates of Orbiting Carbon Observatory (OCO-2) bias corrected data.

This is the latest version of this collection. The DOIs assigned to previous versions, which are no longer available, now direct to this page.

On road emissions data

https://daac.ornl.gov/CMS/guid es/CMS_DARTE_V2.html

DAAC Home > Get Data > NASA Projects > Carbon Monitoring System (CMS) > User guide

DARTE Annual On-road CO2 Emissions on a 1-km Grid, Conterminous USA, V2, 1980-2017

Get Data

Documentation Revision Date: 2019-11-01

Dataset Version: 2

Summary

This data set provides a 38-year, 1-km resolution inventory of annual on-road CO2 emissions for the conterminous United States based on roadwaylevel vehicle traffic data and state-specific emissions factors for multiple vehicle types on urban and rural roads as compiled in the Database of Road Transportation Emissions (DARTE). CO2 emissions from the on-road transportation sector are provided annually for 1980-2017 as a continuous surface at a spatial resolution of 1 km.

DARTE is the first nationally consistent inventory of US on-road CO2 emissions built from bottom-up source activity data and establishes a national benchmark for monitoring, reporting, and verification of emissions that are vital for regulating greenhouse gases. The DARTE data provides insights into how urban areas contribute to climate change.

Bureau of Transportation Statistics

https://www.bts.gov/product/nat ional-transportation-statistics?k eys=miles&field_topic_target_i d=All



Kaggle - can look at competitions and data sets (but pay attention to quality!)

www.kaggle.com

Twitter Climate Change Sentiment Dataset

44k tweets pertaining to climate change



Waste Classification data

This dataset contains 22500 images of organic and recyclable objects

Data Card Code (71) Discussion (0)



Data Card Code (8) Discussion (0)

About Dataset

The collection of this data was funded by a Canada Foundation for Innovation JELF Grant to Chris Bauch, University of Waterloo.

This dataset aggregates tweets pertaining to climate change collected between Apr 27, 2015 and Feb 21, 2018. In total, 43943 tweets were annotated. Each tweet is labelled independently by 3 reviewers. This dataset only contains tweets that all 3 reviewers agreed on (the rest were discarded).

Each tweet is labelled as one of the following classes:

2(News): the tweet links to factual news about climate change

1(Pro): the tweet supports the belief of man-made climate change

 $m{\theta}$ (Neutral: the tweet neither supports nor refutes the belief of man-made climate change

-1(Anti): the tweet does not believe in man-made climate change

The distribution of the data:

Usability ⁽¹⁾ 10.00

License Data files © Original Authors

Expected update frequency

Zindi (can use closed or open competitions)

https://zindi.africa/competitions



Digital Africa Plantation Counting Challenge

Can you create a semi-supervised algorithm to count trees in plantations in Côte d'Ivoire?

Advanced Prediction Computer Vision Agriculture

\$10 000 USD

~1 month to go

76 active participants 268 enrolled

📕 📕 Helping Côte d'Ivoire



Africa Biomass Challenge

Can you predict biomass in cocoa plantations in Côte d'Ivoire?

Advanced Prediction Agriculture

\$10 000 USD

3 months to go

106 active participants 666 enrolled

Helping Côte d'Ivoire



Intron AfriSpeech-200 Automatic Speech Recognition Challenge

Can you create an automatic speech recognition (ASR) model for African accents, for use by doctors?

\$5 000 USD

3 months to go

5 active participants 148 enrolled

Advanced Automatic Speech Recognition Media Health

Driven Data competitions

https://www.drivendata.org/competitions/



DeepCube: Satellite datasets for climate applications

https://deepcube-h2020.eu/technology/earth-sy stem-data-cube/

https://deepcube-h2020.eu/datasets/

Data cube for drought forecasting in Africa

This is a dataset of minicubes suitable for Earth Surface Forecasting, created by the DeepCube research team of the Max Planck Institute for Biogeochemistry. The dataset is focused on Africa and the task is centered around drought impact forecasting.

To learn more about the Africa minicubes please visit <u>https://www.earthnet.tech/docs/ds-africa/</u>

Data cube for the wildfire research community



This dataset has been created by the DeepCube research team of the National Observatory of Athens and is meant to be used to develop models for next day fire hazard forecasting in Greece.

The dataset includes dynamic variables, such as previous day Leaf Area Index, evapotransiration, Land Surface Temperature, meteorological data, fire variables and Fire Weather Index, resampled at daily temporal resolution and Ikm spatial resolution. It also includes static variables, such as roads density, population density and teoperarehic larger.

Open Street Map for general land and road features

https://www.openstreetmap.org

https://wiki.openstreetmap.org/wiki/Environmen tal_OSM



UN agriculture data (crop and livestock, yields, land area, etc)

https://www.fao.org/faostat/en/#data

F	р Г	ood and Agricultu	ire Organizatio	n					powered by Google Q
								s Media	Publications Statistics Partnerships
FAC	OSTAT	Г					<i>ىر</i> ىية	小 中文	English Français Русский Español
俞	Data	Selected Indicators	Compare Data	Definitions and Standards	FAQ			Q	Search an Indicator or Commodity
Dat		/AINS TABLE							
Þ	WHY Production				Þ	Invest	tment 🜔 so	G indicato	r
Þ	▶ ^Q ^{III} Food Security and Nutrition ⁽¹⁾ SDG Indicators					Macro-Economic Indicators			
Þ						▶ ISI Food Value Chain			
▶ ∰*® Trade					▶ ^C Climate Change				
Þ	Prices					► Ø Forestry			
×	• 🕕 Land, Inputs and Sustainability) 🛞 SDG Indicators 🔇			
۲	GCo	st and Afforda	ealthy Diet	▶ [@] World Census of Agriculture					
•	🖏 Po	pulation and E	mployment		▶ അDiscontinued archives and data series				

Remote sensing data sets for crops

https://openreview.net/pdf?id=J tjzUXPEaCu

https://github.com/nasaharvest/ cropharvest

CropHarvest: a global satellite dataset for crop type classification

Gabriel Tseng NASA Harvest gabrieltseng95@gmail.com

Catherine Nakalembe University of Maryland, College Park cnakalem@umd.edu Ivan Zvonkov University of Maryland, College Park izvonkov@umd.edu

Hannah Kerner University of Maryland, College Park hkerner@umd.edu

Abstract

Remote sensing datasets pose a number of interesting challenges to machine learning researchers and practitioners, from domain shift (spatially, semantically and temporally) to highly imbalanced labels. In addition, the outputs of models trained on remote sensing datasets can contribute to positive societal impacts, for example in food security and climate change. However, there are many barriers that limit the accessibility of satellite data to the machine learning community, including a lack of large labeled datasets as well as an understanding of the range of satellite products available, how these products should be processed, and how to manage multi-dimensional geospatial data. To lower these barriers and facilitate the use of satellite datasets by the machine learning community, we present CropHarvest—a satellite dataset of more than 90,000 geographically-diverse samples with agricultural class labels. The data and accompanying python package are available at https://github.com/nasaharvest/cropharvest

Data on what crops are where

https://www.nass.usda.gov/Researc h_and_Science/Cropland/sarsfaqs2. php#Section1_1.0

Research and Science

CropScape and Cropland Data Layers - FAQs

CroplandCROS | CropScape | FAQ | Metadata | National Download | Other CDL Citations

List of CDL codes, class names, and RGB color values: (xls, xlsx)

The new CroplandCROS web app

WASHINGTON, Oct. 28, 2021 – USDA's National Agricultural Statistics Service and Agricultural Research Service have announced enhancements to the CropScape web app, allowing users to more easily conduct area and statistical analysis of planted U.S. commodities. Now known as CroplandCROS (<u>https://croplandcros.scinet.usda.gov/</u>), the geospatial data product hosts the Cropland Data Layer (CDL). The app allows users to geolocate farms and map areas of interest. To aid users, the app features a user guide and instructional videos.

Section 1: CropScape General Information

- 1. What data is hosted on the CropScape website?
- 2. Where can I find instructions on how to use the tools provided through the CropScape website?
- 3. What future updates do you expect for CropScape?
- 4. Who created the Cropland Data Layer and who developed the CropScape web service?
- 5. Why was the Cropland Data Layer created?
- 6. What differences can be expected when comparing CropScape pixel counts and official NASS statistics for counties. <u>ASD</u>, and states?
- 7. Can I download the entire country in one file through CropScape?
- 8. Are color legends available for the Cropland Data Layers?
- 9. Is there a way to view the Cropland Data Layer data on CropScape with just one or two commodities shown at the national, state, district and/or county levels?

Global temperatures

https://www.ncei.noaa.gov/products/l and-based-station/noaa-global-temp

Home / Products / Land-Based Station / NOAAGlobalTemp

NOAAGlobalTemp

The NOAA Merged Land Ocean Global Surface Temperature Analysis (NOAAGlobalTemp, formerly known as <u>MLOST</u> d) combines longterm sea surface (water) temperature (SST) and land surface (air) temperature datasets to create a <u>complete, accurate depiction of global</u> <u>temperature trends</u> d. The dataset is used to support climate monitoring activities such as the Monthly Global Climate Assessment, and also provides input data for a number of climate models.

Extreme weather events

https://github.com/andregraubner/Climat eNet

https://gmd.copernicus.org/articles/14/10 7/2021/

Model description paper | 🞯 🛈

ClimateNet: an expert-labeled open dataset and deep learning architecture for enabling highprecision analyses of extreme weather



08 lan 2021

Prabhat ★, Karthik Kashinath 🗠 ★, Mayur Mudigonda ★, Sol Kim, Lukas Kapp-Schwoerer, Andre Graubner, Ege Karaismailogiu, Leo von Kleist, Thorsten Kurth, Annette Greiner, Ankur Mahesh, Kevin Yang, Colby Lewis, Jiayi Chen, Andrew Lou, Sathyavat Chandran, Ben Toms, Will Chapman, Katherine Dagon, Christine A, Shields, Travis O'Brien, Michael Wehner, and William Collins

Abstract

Identifying, detecting, and localizing extreme weather events is a crucial first step in understanding how they may vary under different climate change scenarios. Pattern recognition tasks such as classification, object detection, and segmentation (i.e., pixellevel classification) have remained challenging problems in the weather and climate sciences. While there exist many empirical heuristics for detecting extreme events, the disparities between the output of these different methods even for a single event are large and often difficult to reconcile. Given the success of deep learning (DL) in tackling similar problems in computer vision, we advocate a DL-based approach. DL, however, works best in the context of supervised learning – when labeled datasets are readily available. Reliable labeled training data for extreme weather and climate events is scarce.

We create "ClimateNet" – an open, community-sourced human-expert-labeled curated dataset that captures tropical cyclones (TCs) and atmospheric rivers (ARs) in high-resolution climate model output from a simulation of a recent historical period. We use the curated ClimateNet dataset to train a state-of-the-art DL model for pixel-level identification – i.e., segmentation – of TCs and ARs. We then apply the trained DL model to historical and climate change scenarios simulated by the Community Atmospheric Model (CAM5.1) and show that the DL model accurately segments the data into TCs, ARs, or "the background" at a pixel level. Further, we show how the segmentation results can be used to conduct spatially and temporally precise analytics by quantifying distributions of extreme precipitation conditioned on event types (TC or AR) at regional scales. The key contribution of this work is that it paves the way for DL-based automated, high-fidelity, and highly precise analytics of climate data using a curated expert-labeled dataset – ClimateNet.

Climate Change AI resource wiki

https://wiki.climatechange.ai/wiki/We lcome_to_the_Climate_Change_Al_ Wiki

Welcome to the Climate Change AI Wiki

Main Page Discussion

Read Edit Edit source View history

This is the approved revision of this page, as well as being the most recent.

The aim of this wiki is to help foster impactful research to tackle climate change, by identifying areas for a useful implementation of machine learning (ML).

The scope of machine learning solutions to address climate change goes far beyond the intersection we address here. Tackling climate change requires cooperation between diverse stakeholders, domain scientists, and action in many forms. Whether you are a machine learning researcher looking to apply your skills to combat climate change, or an early career researcher aiming to have a meaningful impact in your career, a practitioner in one of the domain science areas looking to apply ML to your problem, or for any other reason you are interested in the intersection of climate change and ML, we hope these pages can help inform and facilitate your research!

We welcome your contributions and feedback! This wiki is maintained and moderated by members of CCAI

See guide on contributing to the CCAI Wiki. Feel free to start suggesting changes to any of the following pages!

If you would like to discuss your ideas for additional pages or gain moderator privileges, feel free to reach out to CCAI at wiki@climatechange.ai .

Contents

- 1 Quick start
- 2 Topics by Application Area
- 2.1 Mitigation
- 2.2 Adaptation
- 2.3 Climate science
- 2.4 Tools for Action
- 3 Topics by Cross-cutting Theme

Quick start [edit | edit source]

General Resources page

- Tackling Climate Change with Machine Learning
 <u>review paper or explore its</u> interactive summary
 <u>review paper or explore its</u>
- Explore the Climate Change AI Workshop papers ☑

Want a certain type of data? Check if there is a government agency about it

Or google around

Possible methods

Supervised classification or regression (classic methods, simple artificial neural networks, advanced artificial neural networks)

Computer vision (feature extraction, convolutional neural networks)

Unsupervised methods for representation learning (pca, clustering, artificial neural networks)

Feature importance

Reinforcement learning

More paper examples

https://ai4good.org/fragile-earth-2022/

https://www.climatechange.ai/events#past-events and https://www.climatechange.ai/papers?

e.g., https://www.climatechange.ai/events/neurips2022

https://www.cambridge.org/core/journals/environmental-data-science

https://www.hadr.ai/accepted-papers

Tutorials: <u>https://www.climatechange.ai/tutorials</u>

Advice

Think big (could even result in a publication). This is a growth opportunity.

Choose something meaningful and interesting to you.

Expect things to take awhile and be frustrating, especially as you get started with new data or methods

Document what you are doing for yourself and for your final report

Manage your time! Don't wait til the end of April to get going...

Research Quality

How to avoid machine learning pitfalls: a guide for academic researchers

Michael A. Lones*

Abstract

This document is a concise outline of some of the common mistakes that occur when using machine learning, and what can be done to avoid them. Whilst it should be accessible to anyone with a basic understanding of machine learning techniques, it was originally written for research students, and focuses on issues that are of particular concern within academic research, such as the need to do rigorous comparisons and reach valid conclusions. It covers five stages of the machine learning process: what to do before model building, how to reliably build models, how to robustly evaluate models, how to compare models fairly, and how to report results.

https://arxiv.org/pdf/2108.02497.pdf

Contents

1 2

3

4

5

6

Intro	oduction	1
Befo	pre you start to build models	
2.1	Do take the time to understand your data	
2.2	Don't look at all your data	-
2.3	Do make sure you have enough data	-
2.4	Do talk to domain experts	4
2.5	Do survey the literature	4
2.6	Do think about how your model will be deployed	5
How	v to reliably build models	5
3.1	Don't allow test data to leak into the training process	5
3.2	Do try out a range of different models	(
3.3	Don't use inappropriate models	1
3.4	Do keep up with recent developments in deep learning	8
3.5	Don't assume deep learning will be the best approach	8
3.6	Do optimise your model's hyperparameters	9
3.7	Do be careful where you optimise hyperparameters and select features	9
3.8	Do avoid learning spurious correlations	11
How	v to robustly evaluate models	11
4.1	Do use an appropriate test set	11
4.2	Don't do data augmentation before splitting your data	12
4.3	Do use a validation set	12
4.4	Do evaluate a model multiple times	12
4.5	Do save some data to evaluate your final model instance	14
4.6	Don't use accuracy with imbalanced data sets	14
4.7	Don't ignore temporal dependencies in time series data	15
How	to compare models fairly	16
5.1	Don't assume a bigger number means a better model	10
5.2	Do use statistical tests when comparing models	16
5.3	Do correct for multiple comparisons	17
5.4	Don't always believe results from community benchmarks	17
5.5	Do consider combinations of models	17
How	to report your results	18
6.1	Do be transparent	18
6.2	Do report performance in multiple ways	19
6.3	Don't generalise beyond the data	19
6.4	Do be careful when reporting statistical significance	19
6.5	Do look at your models	20

7 Final thoughts

Writing evaluation

Your writing should be clear and concise, with minimal grammatical errors.

To write your background, find some example papers that may have tackled a similar problem so that you can situate your work relevant to past work [probably should do this before you start!].

For the methods, be clear and complete enough that someone could replicate what you did just from reading your methods section.

Figures should be easy to understand, with clearly labeled axes and not too many different colors or line styles.

Presentation Evaluation

Slides should have clear and informative titles

Figures should be fully visible with labeled axes

Slides should not have too much written text; rather, you should have one or two main points on the slide with one or two major visuals. Elaborate on what is on the slides as you speak.

Speak loud enough for everyone to hear

Go at the right pace throughout (time yourself beforehand!)

Advice: imagine yourself in the audience listening to a project you know nothing about. How would you want the speaker to communicate to you?

Chat GPT

Chat GPT



What are three papers that use machine learning to help building energy efficiency?



Here are three papers that use machine learning to help building energy efficiency:

- "Deep learning-based occupancy prediction for energy-efficient building HVAC control" by X. Liu et al. (2017): This paper proposes a deep learning-based occupancy prediction model for heating, ventilation, and air conditioning (HVAC) control in buildings. The model uses data from various sensors in the building to predict occupancy levels, which is then used to optimize the HVAC system and reduce energy consumption.
- "Building energy consumption prediction using machine learning techniques: A review" by Y. Zhou et al. (2018): This paper provides a comprehensive review of machine learning techniques used for building energy consumption prediction. The authors survey a wide range of methods, including decision trees, neural networks, and support vector machines, and evaluate their performance on various datasets.
- "Integrating machine learning algorithms into a building energy management system for energy-efficient operation" by S. K. Sharma et al. (2020): This paper proposes an energy management system for buildings that integrates machine learning algorithms. The system uses data

Chat GPT



What are three papers that use machine learning to help building energy efficiency?



Here are three papers that use machine learning to help building energy efficiency:

- "Deep learning-based occupancy prediction for energy-efficient building HVAC control" by X. Liu et al. (2017): This paper proposes a deep learning-based occupancy prediction model for heating, ventilation, and air conditioning (HVAC) control in buildings. The model uses data from various sensors in the building to predict occupancy levels, which is then used to optimize the HVAC system and reduce energy consumption.
- "Building energy consumption prediction using machine learning techniques: A review" by Y. Zhou et al. (2018): This paper provides a comprehensive review of machine learning techniques used for building energy consumption prediction. The authors survey a wide range of methods, including decision trees, neural networks, and support vector machines, and evaluate their performance on various datasets.
- "Integrating machine learning algorithms into a building energy management system for energy-efficient operation" by S. K. Sharma et al. (2020): This paper proposes an energy management system for buildings that integrates machine learning algorithms. The system uses data

Nope

Nope

Nope

Deep Learning for Classifying Food Waste

Amin Mazloumian Institute of Embedded Systems Zurich, Switzerland amin.mazloumian@zhaw.ch

Matthias Rosenthal Institute of Embedded Systems Zurich, Switzerland matthias.rosenthal@zhaw.ch

Hans Gelke Institute of Embedded Systems Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich, Switzerland hans.gelke@zhaw.ch

Abstract-One third of food produced in the world for human consumption - approximately 1.3 billion tons - is lost or wasted every year. By classifying food waste of individual consumers and raising awareness of the measures, avoidable food waste can be significantly reduced. In this research, we use deep learning to classify food waste in half a million images captured by cameras installed on top of food waste bins. We specifically designed a deep neural network that classifies food waste for every time food waste is thrown in the waste bins. Our method presents how deep learning networks can be tailored to best learn from available training data.

Index Terms-deep learning, neural networks, machine learning, food waste

The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons, a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed. c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes. In this research, we adopt a deep convoultional neural

network approach for classifying food waste in waste bins [9]. Deep convolutional neural networks are supervised machine

My Research Paper Grace Lindsay

The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons. a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed, c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes. Here, I will test three different neural networks trained to classify food waste

Deep Learning for Classifying Food Waste

Academic Disciplinary offenses include but are not limited to:

Amin Mazloumi: Institute of Embedded S Zurich University of Applie Zurich, Switzerlan amin.mazloumian@zh

Abstract—One third of food 1 consumption – approximately 1 very year. By classifying food w raising awareness of the measu significantly reduced. In this re classify food waste in half a mil installed on top of food waste deep neural network that clas food waste is thrown in the w how deep learning networks ca available training data. Index Terms—deep learning, ing, food waste

- cheating, <u>plagiarism</u>, falsification of data or sources, forgery of academic documents in attempt to defraud;
- destruction, theft, or unauthorized use of laboratory data, or research materials.

The following penalties may be imposed by the faculty for disciplinary infractions:

- Warning
- Disciplinary Probation
- Suspension
- Dismissal

:hnically a difficult / hard to differentiate nsider distinguishing me in a wide variety /aste, e.g. soups and ds, can largely hide st three different

Deep Learning for Classifying Food Waste

Amin Mazloumian Institute of Embedded Systems Zurich, Switzerland amin.mazloumian@zhaw.ch

Matthias Rosenthal Institute of Embedded Systems Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich, Switzerland matthias.rosenthal@zhaw.ch

Hans Gelke Institute of Embedded Systems Zurich, Switzerland hans.gelke@zhaw.ch

Abstract-One third of food produced in the world for human consumption - approximately 1.3 billion tons - is lost or wasted every year. By classifying food waste of individual consumers and raising awareness of the measures, avoidable food waste can be significantly reduced. In this research, we use deep learning to classify food waste in half a million images captured by cameras installed on top of food waste bins. We specifically designed a deep neural network that classifies food waste for every time food waste is thrown in the waste bins. Our method presents how deep learning networks can be tailored to best learn from available training data.

Index Terms-deep learning, neural networks, machine learning, food waste

The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons, a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed. c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes. In this research, we adopt a deep convoultional neural

network approach for classifying food waste in waste bins [9]. Deep convolutional neural networks are supervised machine

My Research Paper Grace Lindsay

According to Mazloumian et al. (2020): "The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons. a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed, c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes." Here, I will test three different neural networks trained to classify food waste

Deep Learning for Classifying Food Waste

Amin Mazloumian Institute of Embedded Systems Zurich, Switzerland amin.mazloumian@zhaw.ch

Matthias Rosenthal Institute of Embedded Systems Zurich, Switzerland matthias.rosenthal@zhaw.ch

Abstract-One third of food produced in the world for human consumption - approximately 1.3 billion tons - is lost or wasted every year. By classifying food waste of individual consumers and raising awareness of the measures, avoidable food waste can be significantly reduced. In this research, we use deep learning to classify food waste in half a million images captured by cameras installed on top of food waste bins. We specifically designed a deep neural network that classifies food waste for every time food waste is thrown in the waste bins. Our method presents how deep learning networks can be tailored to best learn from available training data.

Index Terms-deep learning, neural networks, machine learning, food waste

Hans Gelke Institute of Embedded Systems Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich, Switzerland hans.gelke@zhaw.ch

> The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons, a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed. c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes. In this research, we adopt a deep convoultional neural network approach for classifying food waste in waste bins [9].

Deep convolutional neural networks are supervised machine

My Research Paper Grace Lindsay

According to Mazloumian et al. (2020): "The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons. a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed, c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes." Here, I will test three different neural networks trained to classify food waste

Not plagiarism, but bad writing!

Deep Learning for Classifying Food Waste

Amin Mazloumian Institute of Embedded Systems Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich University of Applied Sciences Zurich, Switzerland amin.mazloumian@zhaw.ch

Matthias Rosenthal Institute of Embedded Systems Zurich, Switzerland matthias.rosenthal@zhaw.ch

Hans Gelke Institute of Embedded Systems Zurich, Switzerland hans.gelke@zhaw.ch

Abstract-One third of food produced in the world for human consumption - approximately 1.3 billion tons - is lost or wasted every year. By classifying food waste of individual consumers and raising awareness of the measures, avoidable food waste can be significantly reduced. In this research, we use deep learning to classify food waste in half a million images captured by cameras installed on top of food waste bins. We specifically designed a deep neural network that classifies food waste for every time food waste is thrown in the waste bins. Our method presents how deep learning networks can be tailored to best learn from available training data.

Index Terms-deep learning, neural networks, machine learning, food waste

The automatic classification of food waste in waste bins is technically a difficult computer vision task for the following reasons. a) It is visually hard to differentiate between edible and not-edible food waste. As an example consider distinguishing between eggs and empty egg shells. b) Same food classes come in a wide variety of textures and colors if cooked or processed. c) Liquid food waste, e.g. soups and stews, and soft food waste, e.g. chopped vegetables and salads, can largely hide and cover visual features of other food classes. In this research, we adopt a deep convoultional neural network approach for classifying food waste in waste bins [9].

Deep convolutional neural networks are supervised machine

My Research Paper Grace Lindsay

According to Mazloumian et al. (2020), visually classifying food waste in bins faces three main challenges: distinguishing edible from non-edible waste, overcoming how diverse single food classes can be, and classifying images with liquid and soft waste that hide other waste classes. Here, I will test three different neural networks trained to classify food waste

Lovely!

Citations and references

\equiv	Google Scholar	Deep Learning for Classifying Food Waste		Q.							
٠	Articles	About 76,900 results (0.18 sec)									
	Any time	Deep learning for classifying food waste		[PDF] arxiv.org							
	Since 2023	A Mazloumian, <u>M Rosenthal</u> , H Gelke - arXiv preprint arXiv:2002.03786, 2	iv:2002.03786, 2020 - arxiv.org								
	Since 2022	food waste can be significantly reduced. In this research, we use deep	learning to class	sify							
	Since 2019	tood waste in bait a million images captured by cameras installed on top of food waste bins									
	Custom range	Save US Cited by 4 Related articles All 3 versions 🔊									
	Sort by relevance	[нтмь] A Novel Model to Detect and Classify Fresh and 🖺	a sea a sea al Possila	- A							
	Sort by date	Reduce Food Waste Using a Deep Learning Technique	×	Cite							
		TB Kumar, D Prashar, G Vaidya, V Kumar Journal of Food, 2022									
	Any type	\ldots to be blamed for the majority of ${\rm food}$ that is ${\rm wasted}.$ Families are the $_{\rm I}$									
	Review articles	waste, and we sought to reduce this by identifying fresh and damaged 1	MLA	Mazloumian, Amin, Matthias Rosenthal, and Hans Gelke.							
		☆ Save 55 Cite Cited by 2 Related articles All 5 versions ≫		arXiv:2002.03786 (2020).							
	include natents										
	include citations	[HTML] An Application of Deep Learning Models to Auton	APA	Mazloumian, A., Rosenthal, M., & Gelke, H. (2020). Deep							
		Classification		arXiv:2002.03786							
	Croate alort	AZ Espinoza - 2019 - search.proquest.com									
	time create alert	are trained and evaluated on a novel food waste dataset to assist in a	Chicago	Mazloumian, Amin, Matthias Rosenthal, and Hans Gelke.							
		of food image classification, rather than the task of food waste clas		"Deep learning for classifying food waste." arXiv preprint							
		☆ Save 切 Cite Related articles All 4 versions		arxiv.2002.03760 (2020).							
			Harvard	Mazloumian, A., Rosenthal, M. and Gelke, H., 2020. Deep							
		[PPF] An efficient classification of kitchen waste using (techniques		learning for classifying food waste. arXiv preprint							
				arxiv:2002.03786.							
		APS Kumar, K Buelaevanzalina - Turkish Journal of Computer and, 20	Vancouver	Mazloumian A, Rosenthal M, Gelke H. Deep learning for							
		removal of knives and chemical waste. • In this project we have to bui		classifying food waste. arXiv preprint arXiv:2002.03786. 2020							
		different types of waste. We used transfer learning models to classify th	E	Feb 6.							
		☆ Save 99 Cite Cited by 2 Related articles All 2 versions ∞		BibTeX EndNote RefMan RefWorks							
		Identification of Food Waste through Object Recognition									
		L Earinella E Fernandes N Michener - 2020 11th JEEE 2020 - josev	nlore ieee ora	[PDF] Receiving							

Put all full citations in the reference section.

Questions?

(would a slack be useful?)