Machine Learning for Science & Society

Contents

Syllabus

- Basic Facts
- Schedule
- Learning Outcomes and Evaluation

Notes

- Spring 2021
- Spring 2022

Reference

- Preparing to present a paper
- Posting Notes

- Spring 2022
- time: MW 3:00-4:15pm
- Professor: Sarah Brown
- course number: CSC 592: Topics in Computer Science
- Credits: 4
- Location: Tyler Hall 108

In this class, we will address the challenges in applying machine learning to scientific research and in high stakes social contexts. On the science side, we will examine the role of ML in research, in particular how it works within knowledge production and how to evaluate ML in line with domain norms. On the social side, we will consider how to ensure ML-based algorithmic decision making systems uphold social values, with a focus on fairness. While these two applications are distinct, many of the challenges translate into common technical problems. Some of the common challenges include:

- missing data
- noisy or missing labels
- multiple objectives

We will look at a range of strategies for identifying and mitigating these problems including:

- robust evaluation
- model inspection
- explanations
- interpretable models

Format

This will be a synchronous course offered in person.

The course will involve:

- reading and evaluating ML research papers
- facilitating and participating in class discussions of the papers
- producing a replication, demo, or illustration of one concept covered for a broader audience
- completing a project using ML in a scientific or social domain
- writing a CS conference style (short & concise) final paper on their project

graduate students are encouraged to do a project related to their research

Prerequisites

To be successful in this class students should have:

- past experience with machine
- basic programming skills
- familiarity with concepts in probability, linear algebra, and calculus that appear in ML

varying skill in these topics is ok, but a general understanding of the basic ideas is important.

Complete this Google form to request a permission number from Professor Brown to enroll in this course. Note that you must be enrolled at URI to take this course and be logged into your URI google account to view that form.

Basic Facts

Meetings

This class will meet on Monday and Wednesday 3-4:15pm in person.

Instructor

Professor Sarah M Brown is an Assistant Professor in Computer Science. Her current research aims to answer the question, “How can machine learning produce AI systems that make fair decisions?”

Office Hours

By appointment, link on Brightspace.

Schedule

We meet in Tyler 108, MW 3-4:15pm.

This course will proceed in three main parts: overview, deep dives, and wrap up.

Structure

Overview

In the first part of the course we will review ML basics, set norms for interaction and complete a survey of the topics that we will cover for the rest of the semester.

In this part of the class, Professor Brown will lead synchronous sessions. Students will be responsible for reading overviews, refreshing background material, and choosing an area for their course project. Students will start with an introductory demo or replication as a mini project.

Deep Dives

During the middle of the course we will spend one week on each topic. There will be 1-3 papers to read each week.

Students will be responsible for presenting papers in class on a rotating basis.

During this time students will have milestones where they need to complete interim steps for their course project. The first milestone will be a proposal that includes the specific products for the remainder of the milestones based on a template.

Conclusion

We will also workshop students’ projects, giving substantive feedback prior to the final submissions.

Final projects will be evaluated through a presentation and paper

Weekly topics

The readings are subject to revision in class up until a presenter is assigned. Topics may also be updated after the first few classes based on student interests and recent publications.

Class	Topic	Reading	Activities
2021-01-24	Introduction	None	introductions, expectation setting
2021-01-26	Probability Review	Model Based ML, chapter 1	reading discussion, setting up
2021-01-31	Setting the Stage	The Scientific Method in the Science of Machine Learning and Value-laden Disciplinary Shifts in Machine Learning	Paper Presentation by Dr. Brown
2021-02-02	Meta issues	Roles for computing in social change	Paper Presentation by Dr. Brown
2021-02-07	Missing Data: Intro strategies	Handling Missing Values when Applying Classification Models & Missing data imputation using statistical and machine learning methods in a real breast cancer problem	Paper discussion led by Emmely & Chan
2021-02-09	Missing data with graphical models and causal reasoning	Graphical Models for Inference with Missing Data	Paper discussion led by Chamudi
2021-02-14	causal and probabilistic missing data	Missing Data as a Causal and Probabilistic Problem	Paper discussion by Lily
2021-02-16	Fairness	Fairml classification chapter and Machine Bias and Gender Shades and Obermeyer	Paper discussion by Damon & Dereck
2021-02-21	Fairness and Causality	FairML Causality chapter	Paper discussion by
2021-02-23	Fairness and experiments	Emprical comparison paper	& empirical setups and scope of work
2021-02-28	Multi-objective & constrained opt	Elastic Net	Paper presentation by
2021-03-02	Multi-objective & constrained opt	A critical review of multi-objective optimization in data mining: a position paper	Paper presentation and discussion by
2021-03-07	Latent Variable Models	Gaussian Mixture Models and Topic Models	Paper presentaiton by
2021-03-09	Latent Variable Models	Indian Buffet Process and Auto-Encoding Variational Bayes	Paper presentation by
2021-03-21	Missing or Noisy labels	Learning with Noisy Labels and Semi Supervised Learning	
2021-03-23	Noisy Labels as a model for Bias	Recovering from biased data: Can fairness constraints improve accuracy and Fair classification with group dependent label noise	
2021-03-28	Interpretable & Explanation Intro	A Survey of Methods for Explaining Black Box Models	Paper Presentation by
2021-03-30	A Case for Interpretability over Explanation	Why are we explaining black box models and Learning Certifiably optimal rule lists for categorical data	Paper Presentation by
2021-04-04	Models for Explanation	Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV) and A unified approach to interpreting model predictions	Paper Presentation by
2021-04-06	Choosing Explanations and using explantions	How can I choose an explainer? An Application-grounded Evaluation of Post-hoc Explanations Actionable Recourse in Linear Classification	Paper Presentation by
2021-04-11	What are the risks of explanations	Model Reconstruction from Model Explanations	Paper Presentation by
2021-04-13	What does Interpretable mean	Towards A Rigorous Science of Interpretable Machine Learning and Towards falsifiable interpretability research	Paper Presentation by
2021-04-18	Project Presentations	projects	Paper Presentation by
2021-04-20	Project Presentations	projects	presentations with peer feedback
2021-04-25	Project Presentations	projects	peer feedback

Class	Topic	Reading	Activities
2021-04-27	Review and Project Reflections	Paper feedback	presentations with revision plans

Table 1 Schedule

Learning Outcomes and Evaluation

This course has goals with respect to the knowledge and research skills.

Evaluation will be with respect to each of the outcomes and based on a level of mastery: general awareness, competency, or mastery.

By the end of the course students will be able to:

- Critique common ways that social or scientific applications of ML require violating ML algorithm assumptions and ways to mitigate or adapt.
- Evaluate ML Research papers for their applicability to scientific and social applications of ML.
- Communicate about ML and its limitations work to varied audiences
- Apply ML to scientific and social data responsibly

Activities

- reading and evaluating ML research papers
- facilitating and participating in class discussions of the papers
- coproducing notes that summarized key points and open questions of papers
- producing a replication, demo, or illustration of one concept covered for a broader audience
- completing a project using ML in a scientific or social domain
- writing a CS conference style (short & concise) final paper on the project

Evaluation

The grading scheme is rooted in achieving the learning outcomes and finalized with a grading contract. Each student will submit a grading contract in the first two weeks and then if all work meets the specification, will earn the contracted grade.

The following describes each activity in the course and the specification for it.

Discussions, Exercises, and Notes

For each topic we cover in class, you should engage fully in the class discussion and practice exercises that are provided if applicable.

To demonstrate engagement you must:

- provide a good faith attempt at any exercises provided
- contribute to the discussion (comments and questions both count)
- contribute to annotated class notes

Presentations

Presenting papers and participating in class will contribute to demonstrating a basic awareness at each of learning objective.

Each class session will be evaluated on if you contribute to discussion or not. This includes both asking questions and answering questions.

Each time you present will be evaluated on specification, your presentation should:

- summarize the key takeaways for the reading(s) in your own words
- summarize key details for understanding to facilitate the discussion
- discussion of strengths and weaknesses of the paper & method
- describe how this paper relates to bigger ideas in the course or your own work

You'll present 2-3 times and you will be expected to improve each time, not to be perfect.

When you present you don't have to have all the answers, you can have open questions.
The goal is that you guide the discussion by doing the above and opening the floor up for questions.

Project

The final project is a chance to dive deeply into one of the course topics. It has the following timeline. Percentages below are of the total grade.

Date	Milestone	Submission format	Evaluation
2022-02-18	Area Selection	Consultation meeting and general questions	feedback only
2021-03-02	Topic Selection	Objectives and scope of work	completion or scope adjustment
2021-03-10	Proposal	Problem statement, lit review, method	specification, with revisions
2021-04-02	Checkin	Consultation meeting and prelim result	completion
2021-04-13	Rough draft	Draft ready for peers to read	feedback only, per paper specs
2021-04-x	Presentation	talk in class	specification
2021-04-26	revision plan	plan for final revision, minor extensions	feedback only, per paper specs
2021-05-x	final paper	final paper submitted for grading	specification
2021-05-x	final reflection	final paper submitted for grading	completion

Proposal Specifications

Submit a 1.5- 2page proposal in the ACM Proceedings format.

Your proposal should include a concise problem statement, a preliminary literature review that situates your project, a description of method(s) you will use to answer your questions in your project, and the expected outcomes of your project.

The proposal will be graded on if it meets the specification or not, but you will be able to revise and resubmit if the first submission does not. To meet specification it must:

- be the right length
- be the right format
- include all sections
- be written clearly
- describe the problem, clearly identifying what the specific goals of your project are
- describe a tractable project
- summarize relevant literature for the problem context
- summarize relevant course-related literature for your project
- describe what you will do in your project
- describe what the end outcome of your project.

Checkin Specifications

- scheduled on time
- at least one dimension of progress from proposal

Presentation Specifications

Your presentation should:

- include an agenda for the talk
- describe the problem

- summarize relevant background
- clearly identify what you did
- describe findings
- include concluding remarks on reflection/possible extensions

Paper Specifications

Your final paper should include a concise problem statement, a complete literature review that situates your project, a description of method(s) used your project, findings, and a discussion or future work section.

For it to meet specification it must:

- be the right length
- be the right format
- include clearly marked sections indicating the required content
- be written clearly
- describe the problem, clearly identifying the specific goals of your project
- summarize relevant literature for the problem context
- summarize relevant course-related literature for your project
- include clear description of what was accomplished
- include a clear summary of results (may include null results/ failed findings)

Translation Mini Project

For this assignment you can choose any topic other than the one your project is for and produce a short demo, illustration, or replication that makes some aspect of the the topic accessible for a broader audience.

For this, you must submit a one paragraph proposal that describes your demo Once that's approved that it will count, you have two weeks to build your demo or replication. The latest your demo may be submitted is at the same time as your final project.

The proposal will be graded on specification and may be resubmitted until successful. Your demo proposal must:

- state the topic from class your demo relates to
- state the format/medium your demo will take:
  - illustration, replication, interactive visualization, etc
- describe the target audience (a particular type of scientists, impacted people, software engineers, layperson, etc)
- describes what your demo will do by answering the relevant questions from the list below:
  - what will a person learn by reading/ using your demo?
  - if it's interactive what will vary? what will be the inputs?
  - what specific result will you replicate?
- describe a demo that is an appropriate scope (not too large or too small)

The demo will be graded on specification and can be revised and resubmitted one time. Your demo must:

- describe a topic accurately
- be accessible to the specified topic model
- meet the description in the proposal

With your demo or after, submit a one paragraph reflection describing what you learned doing this exercise. The reflection will be graded on completion.

Spring 2021

Schedule

We will meet synchronously via Zoom: Tu 5:30-8:15

This course will proceed in three main parts: overview, deep dives, and wrap up.

Structure

Overview

In the first part of the course we will review ML basics, set norms for interaction and complete a survey of the topics that we will cover for the rest of the semester.

In this part of the class, Professor Brown will lead synchronous sessions. Students will be responsible for reading overviews, refreshing background material, and choosing an area for their course project. Students will start with an introductory demo or replication as a mini project.

Deep Dives

During the middle of the course we will spend one week on each topic. There will be 1-3 papers to read each week.

Students will be responsible for presenting papers in class on a rotating basis.

During this time students will have milestones where they need to complete interim steps for their course project. The first milestone will be a proposal that includes the specific products for the remainder of the milestones based on a template.

Conclusion

In the end of the course, we will focus on integrating ideas across multiple topics.

We will also workshop students' projects, giving substantive feedback prior to the final submissions.

Final projects will be evaluated through a presentation and paper

Weekly topics

Class	Topic	Reading	Activities
2021-01-29	Introduction	None	introductions, expectation setting
2021-02-01	Probability Review	Model Based ML, chapter 1	reading discussion, setting
2021-02-03	ML Process & Mutual information preview	Scikit learn getting started,	live coding
2021-02-08	Missing Data: Intro strategies	Handling Missing Values when Applying Classification Models & Missing data imputation using statistical and machine learning methods in a real breast cancer problem	Paper discussion led by Daniel
2021-02-10	Missing data with graphical models and causal reasoning	Graphical Models for Inference with Missing Data & Missing Data as a Causal and Probabilistic Problem	Paper discussion led by Julian
2021-02-15	Current Challenges in Missing data	Handling Missing Data in Decision Trees: A Probabilistic Approach & How to miss data? Reinforcement learning for environments with high observation cost	Paper discussions by Xavier and Zhen
2021-02-17	Current Challenges in Missing data	How to deal with missing data in supervised deep learning	Paper discussion by Madhukara, Replication & testing discussion,
2021-02-22	Fairness	fairml classification chapter and friedler empricial comparison paper	Empirical setup
2021-02-24	Fairness	Reading	preview of lasso and admm constraint to multiobjecitve reformulation
2021-03-01	Multi-objective & constrained opt	Elastic Net	Paper presentation by Daniel, try out elastic net & LASSO in scikit learn
2021-03-03	Multi-objective & constrained opt	A critical review of multi-objective optimization in data mining: a position paper	Paper presentation and discussion by Zhen
2021-03-08	Latent Variable Models	Gaussian Mixture Models and Topic Models	Paper presentaiton by Xavier
2021-03-10	Latent Variable Models	Indian Buffet Process and Auto-Encoding Variational Bayes	Paper presentation by Madhukara
2021-03-15	Missing or Noisy labels	Learning with Noisy Labels and Semi Supervised Learning	Julian and Daniel
2021-03-17	Noisy Labels as a model for Bias	Recovering from biased data: Can fairness constraints improve accuracy and Fair classification with group dependent label noise	Zhen
2021-03-22	Interpretable & Explanation Intro	A Survey of Methods for Explaining Black Box Models	Xavier
2021-03-24	A Case for Interpretability over Explanation	Why are we explaining black box models and Learning Certifiably optimal rule lists for categorical data	Madhukara
2021-03-29	Models for Explanation	Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV) and A unified approach to interpreting model predictions	Zhen
2021-03-31	Choosing Explanations and using explantions	How can I choose an explainer? An Application-grounded Evaluation of Post-hoc Explanations Actionable Recourse in Linear Classification	Daniel
2021-04-05	What are the risks of explanations	Model Reconstruction from Model Explanations	Xavier
2021-04-07	What does Interpretable mean	Towards A Rigorous Science of Interpretable Machine Learning and Towards falsifiable interpretability research	Madhukara
2021-04-12	Meta issues	The Scientific Method in the Science of Machine Learning and Value-laden Disciplinary Shifts in Machine Learning	Sarah

Class	Topic	Reading	Activities
2021-04-13	Meta issues	Roles for computing in social change	Sarah
2021-04-19	Project Presentations	projects	presentations with peer feedback
2021-04-21	Project Presentations	projects	peer feedback
2021-04-26	Review and Project Reflections	Paper feedback	presentations with revision plans

Table 2 Schedule

Class 1: Introductions

Introductions & Goals

Course Admin

- Brightspace
- Zoom
- Google docs or markdown in the future?
- Website

Learning outcomes

knowledge research

- identify common problems and solutions in scientific application of ML
- identify common challenges and solutions for social applications: fairness,
- implement and extend research papers

Activities

- reading and evaluating ML research papers
- facilitating and participating in class discussions of the papers
- producing a replication, demo, or illustration of one concept covered for a broader audience
- completing a project using ML in a scientific or social domain
- reflect on methodologies used in this type of research
- writing a CS conference style (short & concise) final paper on their project

Model Based ML and this course

https://www.mbmlbook.com/toc.html

- missing data
- noisy or missing labels
- multiple objectives

We will look at a range of strategies for identifying and mitigating these problems including:

- robust evaluation
- model inspection
- explanations
- interpretable models

ML and Probability Review

admin

- collaborative notes
- brightspace will be updated later this week
- grading details by Wed
- environment for coding demos

More formalism

- model
- prediction algo
- cost function
- objective

Probability

- sample distros

Practical Application of ML & Pipelines

Class 3: ML Pipelines

Goals when using ML

1. Understand about the data (data science/ actual science) probability more statistics, maybe fit another examine model parameters, inspect them
2. understanding about Naive bayes fit different data varies
3. claims about the learning algorithm run multiple algorithms on the same data possibly multiple data

Basic setup

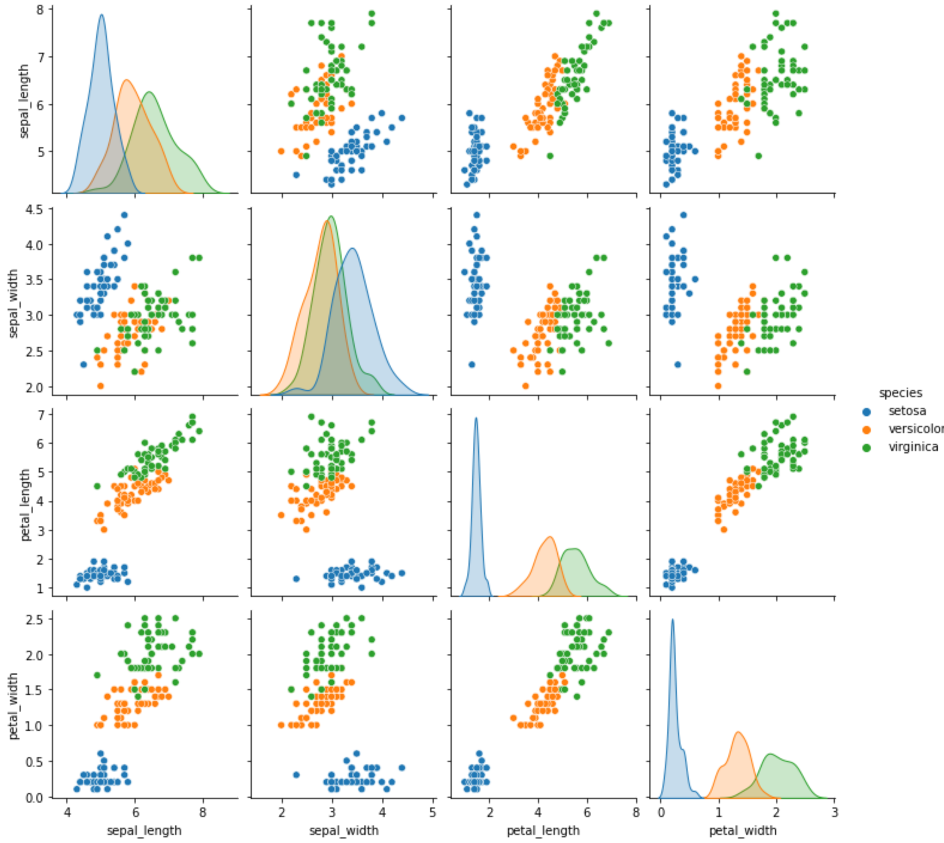
1. test train
2. training parameters
3. estimator objects
4. fit model parameters
5. metrics
6. cross validation

```python
import pandas as pd
import seaborn as sns
import numpy as np
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import confusion_matrix, classification_report
from sklearn import datasets
```

```python
iris_df = sns.load_dataset('iris')
```

```python
sns.pairplot(iris_df, hue='species')
```

```
<seaborn.axisgrid.PairGrid at 0x7f888528e650>
```




```
X,y = datasets.load_iris(return_X_y=True)
```

```
X.shape
```

```
(150, 4)
```

```
y.shape
```

```
(150,)
```

```
X_train, X_test, y_train, y_test = train_test_split(X,y,)
```

```
gnb = GaussianNB()
```

```
gnb.__dict__
```

```
{'priors': None, 'var_smoothing': 1e-09}
```

```
gnb.fit(X_train,y_train)
```

```
GaussianNB()
```

```
gnb.__dict__
```

```
{'priors': None,
 'var_smoothing': 1e-09,
 'classes_': array([0, 1, 2]),
 'n_features_in_': 4,
 'epsilon_': 3.101274713010204e-09,
 'theta_': array([[4.95128205, 3.38717949, 1.47179487, 0.23846154],
        [5.96111111, 2.76388889, 4.33055556, 1.33888889],
        [6.54594595, 2.97297297, 5.4972973 , 2.03243243]]),
 'var_': array([[0.10660092, 0.09701512, 0.03023011, 0.01159764],
        [0.2479321 , 0.10619599, 0.19545525, 0.04070988],
        [0.44464573, 0.11440468, 0.31647918, 0.05894814]]),
 'class_count_': array([39., 36., 37.]),
 'class_prior_': array([0.34821429, 0.32142857, 0.33035714])}
```

```
X_test[0]
```

```
array([6.2, 2.9, 4.3, 1.3])
```

```
y_pred = gnb.predict(X_test)
```

```
y_pred[:5]
```

```
array([1, 0, 0, 0, 1])
```

```
y_test[:5]
```

```
array([1, 0, 0, 0, 1])
```

```
confusion_matrix(y_test, y_pred)
```

```
array([[11,  0,  0],
       [ 0, 14,  0],
       [ 0,  2, 11]])
```

```
gnb.score(X_test,y_test)
```

```
0.9473684210526315
```

```
gnb2 = GaussianNB(priors=[.5,.25,.25])

gnb2_cv_scores = cross_val_score(gnb2,X_train,y_train)
```

```
np.mean(gnb2_cv_scores)
```

```
0.9644268774703558
```

```
gnb_cv_scores = cross_val_score(gnb,X_train,y_train)
```

```
np.mean(gnb_cv_scores)
```

```
0.9644268774703558
```

```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        11
           1       0.88      1.00      0.93        14
           2       1.00      0.85      0.92        13

    accuracy                           0.95        38
   macro avg       0.96      0.95      0.95        38
weighted avg       0.95      0.95      0.95        38
```

```
gnb.predict_proba(X_test)
```

```
array([[5.62306973e-082, 9.99434685e-001, 5.65314849e-004],
       [1.00000000e+000, 1.43668612e-016, 1.15842362e-024],
       [1.00000000e+000, 5.26467677e-020, 1.27341782e-027],
       [1.00000000e+000, 4.07435879e-019, 1.69981016e-026],
       [8.15674694e-034, 9.99999647e-001, 3.53061292e-007],
       [2.13353323e-268, 6.31077883e-010, 9.99999999e-001],
       [5.88357634e-083, 9.99170257e-001, 8.29743094e-004],
       [2.39342456e-092, 9.88212675e-001, 1.17873253e-002],
       [2.68411939e-088, 9.99648212e-001, 3.51788359e-004],
       [5.23628297e-187, 7.24934155e-004, 9.99275066e-001],
       [1.00000000e+000, 9.92773071e-020, 2.09315072e-026],
       [7.85922469e-128, 8.18994832e-001, 1.81005168e-001],
       [1.14630713e-095, 9.92564431e-001, 7.43556919e-003],
       [1.62850059e-266, 2.70939488e-010, 1.00000000e+000],
       [2.82050236e-096, 9.91801078e-001, 8.19892243e-003],
       [2.71241839e-188, 1.53410727e-006, 9.99998466e-001],
       [1.00000000e+000, 4.37467358e-015, 4.43647280e-022],
       [6.28937490e-166, 3.17384338e-003, 9.96826157e-001],
       [7.26782070e-150, 6.53653517e-001, 3.46346483e-001],
       [5.71459159e-085, 9.96429055e-001, 3.57094465e-003],
       [1.92110243e-178, 1.13979019e-003, 9.98860210e-001],
       [1.00000000e+000, 7.58672145e-017, 5.08119127e-023],
       [2.06504121e-192, 8.23550048e-007, 9.99999176e-001],
       [1.00000000e+000, 3.02569784e-019, 6.92814464e-027],
       [1.00000000e+000, 1.16663412e-018, 2.16449638e-026],
       [1.00000000e+000, 6.87221015e-019, 1.62832214e-026],
       [1.85389920e-226, 1.41739900e-008, 9.99999986e-001],
       [3.15469536e-029, 9.99999340e-001, 6.59949380e-007],
       [8.05787768e-199, 3.77621320e-006, 9.99996224e-001],
       [2.29747280e-054, 9.99997362e-001, 2.63768379e-006],
       [4.43832939e-054, 9.99965022e-001, 3.49779686e-005],
       [1.00000000e+000, 1.62240502e-018, 6.71038620e-026],
       [3.91612148e-070, 9.99853945e-001, 1.46054568e-004],
       [4.15299064e-062, 9.99994904e-001, 5.09626789e-006],
       [3.88004465e-185, 1.96916583e-006, 9.99998031e-001],
       [6.12002977e-110, 8.73592735e-001, 1.26407265e-001],
       [7.13769082e-229, 1.05597936e-009, 9.99999999e-001],
       [1.00000000e+000, 9.18560427e-018, 1.95348576e-024]])
```

Class 4: Missing Data: Basic techniques

Evaluation of missing data at training

- mulitple imputation
- ML based was better than imputation which is better than dropping samples
- example datasets: 45% of patients have at least 1 missing value

Imputation

- Mean imputation:
    - insert the mean based onthe other values
- Hot deck
    - mean-like with similarity
- Multiple inputation
    - 3 diff ways

Imputation ML

- MLP
  - fully connected
- Self organization
  - competitive learning
  - NN on modle of nodes in 2d grid,
- KNN
  - select closest complete case to impute values from
  - expensive for large datasets due to need to search everywhere for each missing value

Testing

- Train NN based on data imputed with each technie

Conclusions:

- in general, any imputation was better than deletion
- ML based performed better

Discussion & Questions

- interesting that even simple methods provide improvement
- SOM is sort of unclear how does that work?
- Review of MLP and [sigmoid]

Handling missing values At application time

- reduced models vs imputation.
- broad approach
- 15 common datasets

Techniques:

- Discard
- Acquire missing values
- Imputation
  - predictive value imputation
  - distribution based
  - unique values
- Reduced Feature Models
  - retrain for different feature models

Feature imputability impacts the distribution or predictive type of imputation

More complex model

- decision tree with bagging
- again, reduced model is the best strategy

Hybrid Models for efficient prediction

- reduced models
- a hybrid is a complete model with stored subset for most common missing features
- Reduced feature enseble
  - N models for N features
  - each one is missing one feature
  - average these together for final prediction
  - substantial reduction in when there is a single feature is missing

- combine with imputation for multiple features
- relative accuracy is better than imputation

General takeaways

- reduced models vs imputatation is a large improvement
- this is sort of an imputation

Weaknesses

- Didn't check unique value imputation
- MCAR
- focused on

Overall Discussion

- How might the two problems interact?
    - if missing data at both train and prediction…
    - train using missing data without imputation for training the separate models
- Questions on these ideas
- What additional things might you need to consider when choosing one?
    - feature imputability at training
- what to do with time series data
- How to check if missing CAR?
    - look at collection technique
    -
- what do to with varying data per person
    - LSTM for time series data
    - hierarchichal modeling other wise
    - example of hierarchical with time series also

For Wednesday

1. Graphical Models for Inference with Missing Data
2. Missing Data as a Causal and Probabilistic Problem

Missing Data 2:

Graph theory foundation

- A DAG
- shapes are nodes
    - nodes generally represent a random variable
- nodes are connected with edges
    - edges may be directed (with an arrow)
- path is a sequence of edges
- a cycle is a path that returns to a given node twice
- we will focus on acyclic graphs
- directed edges connect parent nodes to child nodes (follwoign the arrow)
- Why graphs: useful representation of joint distribtions
- d-connected: two nodes are d-connected if there is a connected path without a collider
- d-separation: independent through a collider
- collider is when arrows flip

Missingness graphs

- x,y are variables
- Y^* is a proxy for y
- R_y : causal mechanism for missingness of y*

Recoverabiilty for MCAR

Discussion

- proxy
- example with ocean data temp sensor, cloud cover images

For next week

Choose one: https://artemiss-workshop.github.io/#program

Information Theoretic Approaches for Testing Missingness in Predictive Models https://openreview.net/forum?id=6Y05VJfGlFM

Missing Data 3

Handling Missing Data in Decision: A probabilistic approach

key ideas

- A decision tree's structure and notation
- Review of imputation
  - Predictive value imputation
    - mean, median or mode
    - make assumption that features are indpendent
    - surrogate splits, partition data using another feature to
- XG Boost

Expected Predictions:

- impute all possible completions as once to avoid strong dist assumptions
- consistent for MCAR and MAR
- expensive, but density can help reduce
- tractably compute the exact expected predictions
- loss minimization

Experiments

- for a single dataset, outperforms in general

Discussion

- generally easier
- given single dataset, of results, how much do we trust this?
- what does this provide as an advantage
- NP hard

How to miss data?: Reinformcent leanring for environgments iwith high obseration cost

Key points

Reinforcement learning

- cost associated with making accurate observations
- goal directed
- RL agent tries to

Problem setting:

- \(o \sim p_0(o_t |s_t; \beta)\)
- beta is accuracy og obs
- r is old reward

Scenario A:

- observed cangle vs

Big picture: manipulating how the data collection

Discussion

- survivorship bias?
- right left imbalance for figure 3
- simple pendulum example helped overcome the background lacking
- figures

General

Try writing out a missingness graph for a problem of choice, some scenario where you imagine there would be missing data, or an example dataset that you can find.

MIssing data

supervised

Background

- Hadamard
-

Readings for next week:

http://sorelle.friedler.net/papers/fairness_comparison_fat19.pdf https://fairmlbook.org/

- introduction and classification chpaters (1 and 2)

Elastic Net

1. OLS can overfit
2. ridge helpw sith over fitting, but not variable selection
3. lasso helps reduce the dimensionality

- p>n lasso saturates at n variables
- lasso predicts 1 of correlated variables at random
- ridge is better in correlated case

Multiobjective

Latent Variable Models: GMM & Topic Models

Gaussian Mixture models

- key points:
    - model versus algorithm
- algorithm:
    - initialize
    - Estep
    - Mstep
    - until convergence:
        - parameters stop changing, assignments stop changing
- Covariance types:
    - covers weakness in kmeans

Topic Models

- corpora: collection of documents
- text modeling, was classically binary matrices
    - also tf-idf
    - useful for dicsriminating docuemtns,
    - lacks meaning
- pLSI: probabilistic, latent semantic indexing
    - reminiscnet of GMM
    - assumes exchangeability
    - mixutre components

Latent Variable Models:

Semi-supervised learning and noisy labels

Key questions:

How do these reltae?

Noisy labels as a Bias model

Fairness constraints for recovering from biased labels

Blum & Stengl

Considers 3 cases:

- more errors in the disadvantaged group than the advantaged group
- fewer positive examples of the disadvantaged group
- both

Comparison of Fairness Interventions

Paper discussion

Spring 2022

Notes will be added after the semester starts.

Overview

Course Info

- graduate course, focused on research adjacent skills
- topic is how to use ML safely and reliably in the context of scientific discovery or social applications
- Classes will mostly be discussion
- We'll rotate leading the discussion
- we'll rotate note taking

Intros and Topics of Interest

- how to understand bias and what can be done, multiple dimension to explore
- more about reading and writing papers
- more skill in reading research papers
- missing data, incomplete problems
- HCI
- breadth, more research
- ML
- eg (pain area)
- noisy data
- natural disaster evacuation plan
- incomplete data
- NLP

Overview of Course Topics

- COMPAS Example
- disparate treatment/impact
- medical
-

Prepare for the next class

Prepare for Wednesday:

Model Based ML: https://mbmlbook.com/toc.html

Read: Chapter 1 & the Interlude on the ML life cycle Skim the intro to two application chapters Be prepared to compare this view of ML to how you've learned int (or other CS topics previously)

Read: https://web.stanford.edu/class/ee384m/Handouts/HowtoReadPaper.pdf

Be prepared to ask questions about how to prepare for presenting a paper in class

Create or make sure you can log into GitHub Account

2022-01-26

Lead Scribe: Lily

Admin

- sorry about notes
- private github repo –> Spring 2022
- grading contract FYI

    - Will be given further instructions on ways to achieve an ‘A’ or ‘B’
        - To get a ‘B’ you will only need to complete the paper and presentation
        - To get an ‘A’ you will implement a project (translation)
    - Paper and presentation will be assigned
        - Paper –> CS Conference Style
        - Draft due: last day before the presentation, will be posted for the class to review

Opening Question

What kind of data are you most in working with?

- Class response:
    - GIS data
    - Linguistic data (tweets, reddit posts)
    - Numerical data (tabular)
    - Video/Image
    - Time series
    - EHR/Medical related data
    - NLP
    - tabular/survey

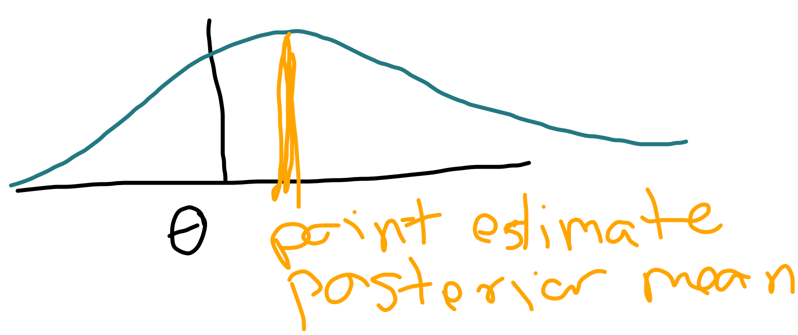
How to Read a Paper

Model Based ML

- Discrete probabilities (distributions introduced in murder mystery chapter)
- Bernoulli
- Priors (probablistic guess about a random variable)
    - Are useful for working with less data to create strong inferences
        - Working with things when not a lot of data is available
    - Assumptions, expressed in a probability distribution
- Posterior
    - Inference given regularizer: Likelihood…
    - Most common posterior probability distribution we’re doing: Probability of parameters given data
- Point Estimate
    - This are the single values produced after training (weights)

$$P(\text{parameters} \mid \text{data})$$

    - Posterior mean



- Most of the probability distributions we’ll use belong to the exponential family
    - https://en.wikipedia.org/wiki/Exponential_family

- Conditional Probability
  - One for each value of the conditioning variable
  - (e.g.) Murder mystery –> murderer variable can be Grey or Auburn

$$P(\text{weapon}) = \sum_{\text{murderer}} P(\text{weapon}, \text{mu})$$

$$P(w=d) = .03 + .56 = .59$$

$$P(w=r) = .27 + .14 = .41$$

- Marginal probability
  - (Section 1.2 – A Model of Murder)
  - "Probability of one event in the presence of all (or subset) outcomes of the other random variable…" (https://machinelearningmastery.com/joint-marginal-and-conditional-probability-for-machine-learning/)
- Maximum Likelihood Estimation
  - Assume a distribution, our goal will be to find the theta (parameter)
  - Maximizing, find parameters that will give us the highest probability (finding the one–parameter–that fits best)
- elicitation - an interdisciplinary field in statistics and psychology; study of how to get an expert's distribution for how likely an event is to occur.

Prepare for next class

- Order of the weekly topics may change
- Dr. Brown will present next week, but we'll start rotating the following week
- There are (2) readings, bring questions and prepare

Learning & Evaluation

- Read through the whole Learning and Evaluation Page after I post a notification to, there are some fixes to be made
- Bring Questions to class next week
- Be ready to work on your grading contract

Reading

The Scientific Method in the Science of Machine Learning and Value-laden Disciplinary Shifts in Machine Learning

Scientific Method & Philosphy of ML

Lead Scribe: Derek Jacobs

admin

- grading contract will be posted in time for Wednesday
- notes & posting example
- can leave out admin in notes; that material wil mostly be other places

Opening Notes

- set the stage for how we think about other work and setting up your projects

Scientific Method in the Science of Machine Learning

[paper]

Introduction

- ML is having a hard time explaining some results
- Replications are failing

Scientific Method

> Starting from the assumption that there exists accessible ground truth, the scientific method is a systematic framework for experimentation that allows researchers to make objective statements about phenomena and gain knowledge of the fundamental workings of a system under investigation.
>
> —Jessica Zosa Forde and Michela Pagnini

- process and a social contract
  - ML might not be systematic
    - Control the randomness of our algorithms
  - Procedures/Trust allow comparisons of results
- assumes a ground truth
  - Things can go wrong if there simply isn't a ground truth
  - Especially when using ML in social context
- Hypotheses
  - A formed "guess"
  - In CS, the learning process is as follows
    - Here's a problem (that has an answer)
    - Write code to solve it to specifications
  - Hypotheses can be falsified through statistical and other analysis
    - You can prove the opposite is not true, but challenging to prove truth

you can express hypotheses as priors, but that prior wouldnt be the same in the rest of ML lit

> **Discuss**
>
> When might this not apply?

> At the base of scientific research lies the notion that an experimental outcome is a random variable, and that appropriate statistical machinery must be employed to estimate the properties of its distribution.

- Treat your accuracies as a random variable as part of an experiment and you're experimenting around those
- In stats people chase having p values <= 0.05 and so we have a reproducibility crisis
- [Reproducibility Study]
  - < 40% replicated the original results
- there are some problems with NHST but there are alternatives that are still rigorous
  - [for design research]
  - [in psych]
  - [bayesian in hci]
    - Bayesian
      - Priors allowed

        - More broadly defined
        - Everything we do is influenced by our thoughts and so there's some subjectivity
        - Interpreting results in terms of previous results
    - Frequentist
        - No Priors
        - Probabilities strictly refer to events like fair coin flips (100 flips we expect 50 heads 50 tails)
        - Knowledge does not accrue

Case study

HEP:

- proposed thoery
- null hypothesis
- cafeull accounting
- model building and hypothesis testing phases
- parametric models derived from first prcinicples
- statistical test is constructed

ML Analogy:

- suspect new activation
- formulate quantitative hypothesis
    - How will it work, what'll it do, how much will it improve
    - And behavior of how it'll change
- run experiments
    - Record outcomes from base models (no intervention)
- This is a statistical model of your experiment, not an actual model
    - Dataset, optimizer, init, hyperparameters, etc are just noise in the question of "does the activation function work"
    - Or make things specific (improve results with a specific dataset)

Recommendations

1. formulate hypotheses first 2. statistical testing 3. operate in controlled, reproducible, and verifiable settings 4. negative result workshops - pregistration workshop at neurips 2021 not an author as speaker - If your experimental plan is good, results are published regardless of positive or negative case - new journal

Value Laden Shifts in ML

paper @brownsarahm reminder from lily **talk on incorporating ethics into teaching data structs - environmental cost of algorithms**

- Looking at how different values influence what is being done in ML
- disciplinary shifts are not objective, but value laden
- Model Types
    - Different categories of models typically used for specific purposes
    - The structure of what our learning algorithm outputs
    - e.g. CNN for image processing, Linear Models, SVMs, etc

Model types as organizing

- many researchers self-organize in types; eg for revieing, workshops etc
    - How do we organize them in a package like sklearn, but also "who knows who works on this"
- commitment is fueled by exemplars
- has down stream effects:
    - guide research agenda and problem selection
        - How model types influence problem selection
            - Different types are tailored to different problems
            - Most popular model means most funding

            - More funding on specific models drives improvements to that area specifically and other problems get lost
    - constrain search for solutions
    - prerequisites, eg deep learning and data volume (fig 1) and compute power (fig2)
        - fig1
            - Depending on data amount, we may favor one model over the other, and that in turn is researched more
        - fig2
            - Shows over time how my computer power is used in training AI
            - Recently exponential growth
- model types have parallels but important differences in philosophical scientific organzing principles
    - decreases theory development
    - Kuhn
        - The organizing paradigm defines what questions are valid
        - For example
            - If we only have earth water fire ether, we can't ask what molecules do
- When committed to model types…
    - The whole industry shifts towards it
    - Like NVidia GPUs, computer architectures, etc

Model type is self-reinforcing

- comparing them is influenced by the model type points above (problems, prerequisites)

Comparison between types is value laden

- Applied to ImageNet
    - Benchmark image classification problem
    - Until 2011, best error rate was 25% (no deep learning)
    - 2012 - AlexNet reached 16% (deep learning)
    - By 2016 ImageNet is basicaclly "done" due to all extremely high accuracies (97% vs 97.1%)
- This doesn't mean the problem of image classification is solved
- prerequisites
    - compute
        - Who has access
    - data
        - Who has access
        - What sets are created/curated
- Evaluation criteria
    - in theories: as a whole, internal consistent, predictive, etc

> Evaluating theories based on their theoretical virtues is a value-laden activity when theoretical virtues are carriers of values.
>
> —Milli and Dotan

- eg consistency requires keeping the bad from the old in new
- eg: metrics and discrimination
    - Something that was sexist previously is still now

Conclusion

> A related question inspired by these issues is who should make decisions in what values are furthered? Who gets to have a voice? In talking about selection of problems in science, Kitcher (2011) argues that all sides should have a say, including laypersons [71]. A question for machine learning is: is the same true for machine learning? Who should have a say about which criteria are important in evaluating model-types? That is itself another value-laden question.

Overall Paper thoughts

> **Discuss**
>
> General thoughts?

- Think about things that're interesting, confusing, etc (for future reference)
- Take care when conducting ML
    - Feed in data, hope for the best without considering the "why's"

> **Discuss**
>
> How is this different from how you've thougth about CS before?

- There's more to consider before actually coding
- Vast social influences in things as simple as "can i even get this dataset"
- We focus on what works now instead of what hasn't worked

> **Discuss**
>
> How might the value-laden points about theory development relate the the scientific method points

- Two pillars of hypothesis + statistical testing

Meta points

- (science) scientific mehtod one is a workshop paper (approx length of your project papers)
- this is a position paper (it's not about new experimental results as much as the classic research arguing a position apper)

next class

- volunteer: Emmely
    - [Roles for computing in social change](https://dl.acm.org/doi/abs/10.1145/3351095.3372871
    - See course site for notes on expectations during presentations

---

Roles for Computing in Social Change

Lead Scribe: Damon Coffey

Roles for Computing in Social Change -concerns about fairness, bias and accountability in the field

Introduction:

- high stakes decision making algorithms have potential to predict outcomes more accurately
- cs has generally failed to target the correct point of intervention
- ex: intervention at the selection phase in an employment context could prevent a hostile work place

Computing as a Diagnostic

- computing can help us measure social problems and diagnose how they manifest in tech systems
- computing cannot solve issues on its own
- Diagnostics work can be valuable
  - highlight tech dimensions of social problems
- misinformation can negatively affect marginalized populations more ex: search engines displaying low quality health information
- not presented as solutions, rather as tools to document practices
  - not to confuse diagnostics with treatment
  - computing is not unique in helping diagnose social problems
    - sociology, etc..
  - certain tools can be treated as certainity for every situation, which is not the case

Computing as a Formalizer

- computing requires explicit specification of inputs and goals
- these inputs and goals can be affected by transperency, accountability and stake holder participation
  - need to be precise ex: risk assessment: debate over how to formalize pretrial risk, if and how to use these instruments
- not all data is easy to quantify
- may press people to rely on measures that are incorrect

Computing as Rebuttal

- computing can clarify the limits of technical interventions and of policies promised on them
- limits of computing can drive people to reject computational approaches
- ex: using an algorithm to determine an immigrant's societal worth, not good. Should seek a different method rather than forcing a technological one
- need to understand what algorithms are actually capable of, instead of forcing it on everything
  - need to show what an algorithm CANT do (prove limits)
- prediction algorithms for risk assessment
- computational research on fairness is built on discrimination law
- Risks
  - proclomations of what a computational tool is incapable of may focus on improving tool even if it is not possible

Computing as a Synecdoche

- computing can foreground long standing social problems in a new way
- Eubank's core concern: computing is just one mechanism through which longstanding poverty policy is manifested
- Automated systems can divert poor people from the resources they need
- computing can help bring attention to old problems, however
- synecdochal focus on computing must walk a pragmatic line between over emphasis on tech aspects and recognition of the work tech actually does
- need to find a balance between the two and develop better systems with more emphasis on social issues

Preparing to present a paper

Questions that will help organize your preparation, but may apply variably to different readings:

- What is the key question that drove the research?
- What is the main finding?
- What is the model assumed in the paper?
- Did they include experimental results? If so:
  - do the experiments support the claims?
  - what additional experiments would help make the result make more sense?
  - how broad are the experiments, are the context-specific or general?
- Is there an analytical result? if so:
  - do the conditions for the proof make sense?
  - are they realistic?
  - what questions do you have about the proof?

You may plan to use slides if that make you more comfortable or you can show the paper itself. You may also show other materials if appropriate and you can seed the day’s notes.

Posting Notes

First time:

1. Go to the notes page of the repository
2. Click add a file, choose create a new file
3. Add your notes
4. At the bottom choose “propose changes”
5. If applicable, navigate to your fork, to the branch you made to add additional files (eg images)
6. Open a pull request from your fork/branch to the the course repo/main.

> **Tip**
>
> These are a rough outline, if you need help, definitely feel free to ask.
> Once you do it, feel free to add more detail.

By Sarah M Brown