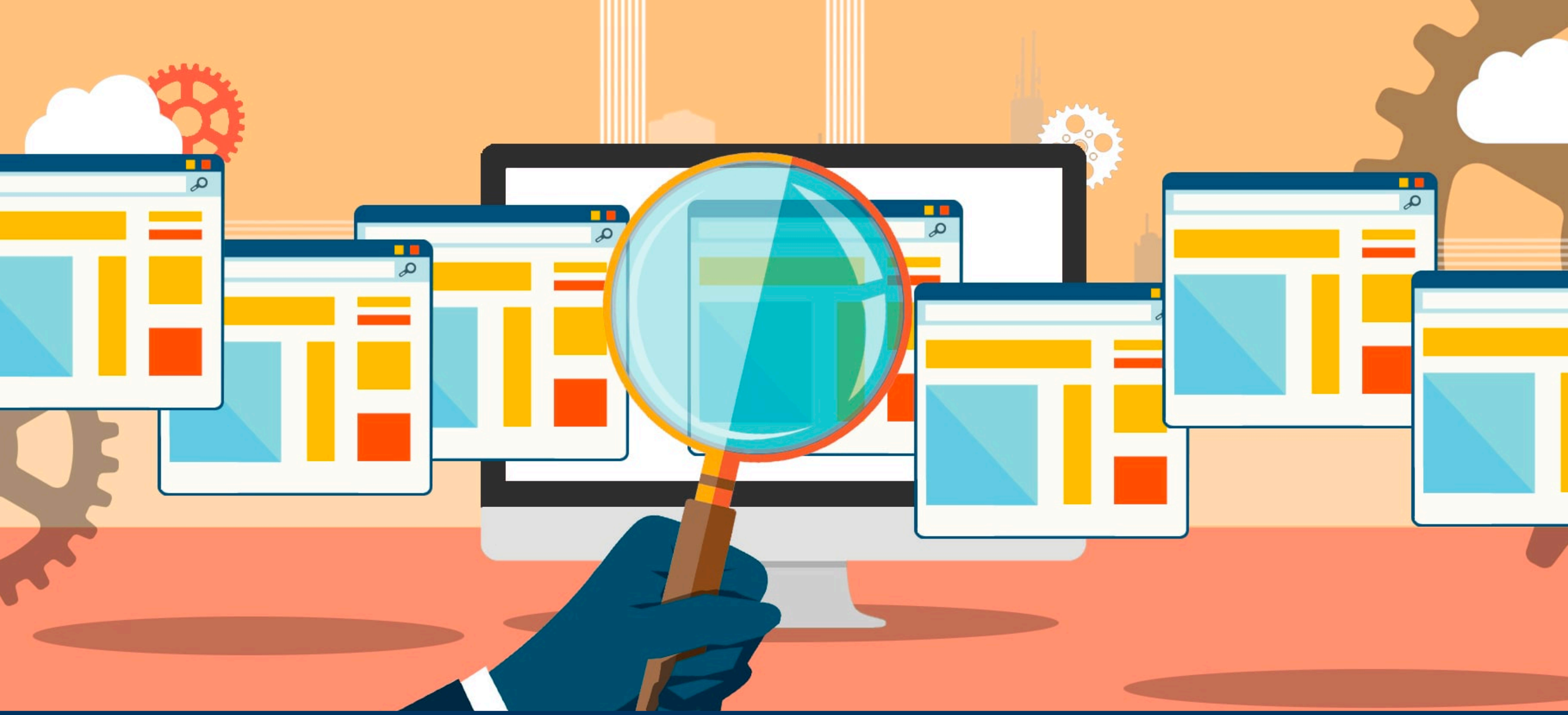
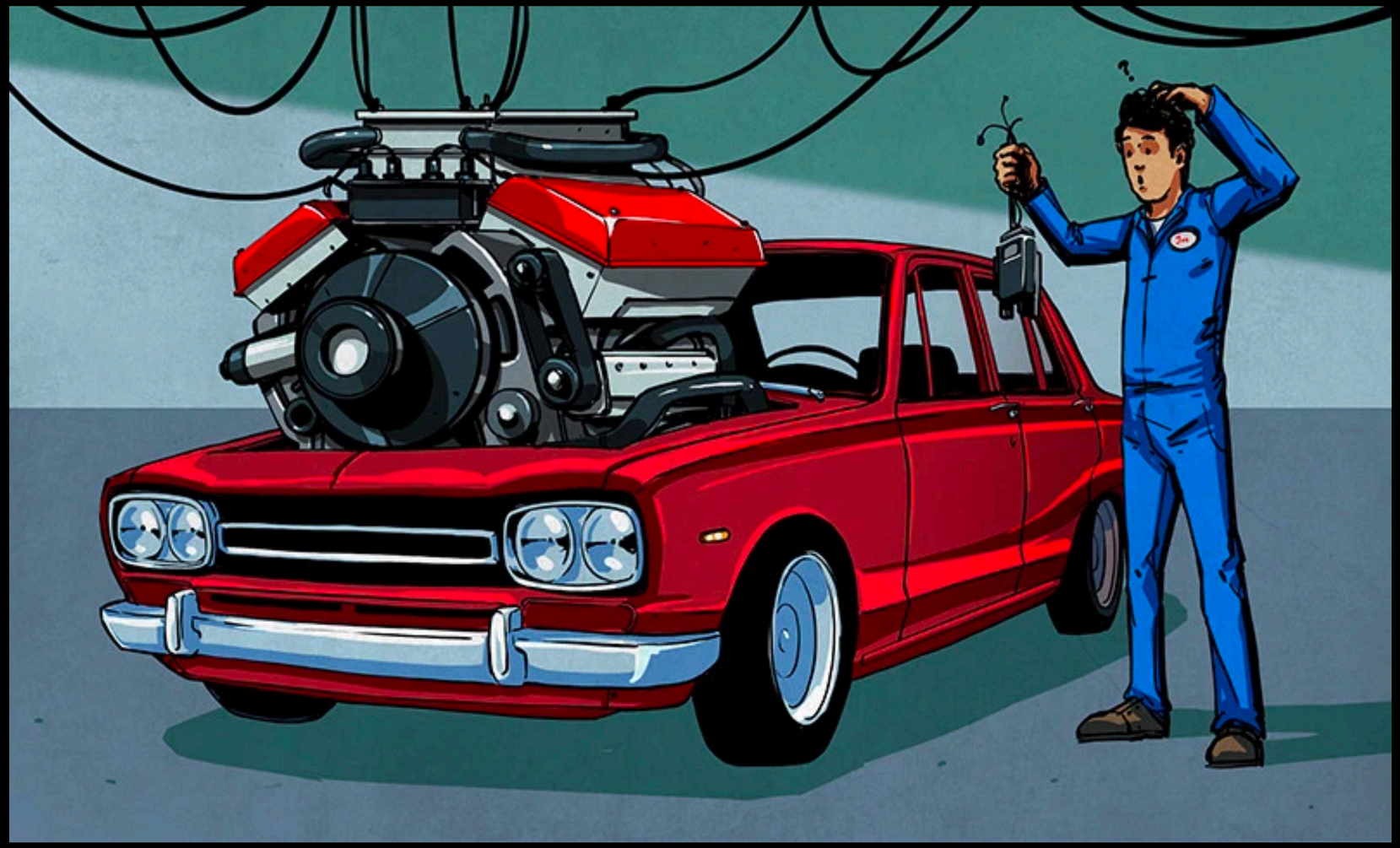


Evaluation in Supervised Learning

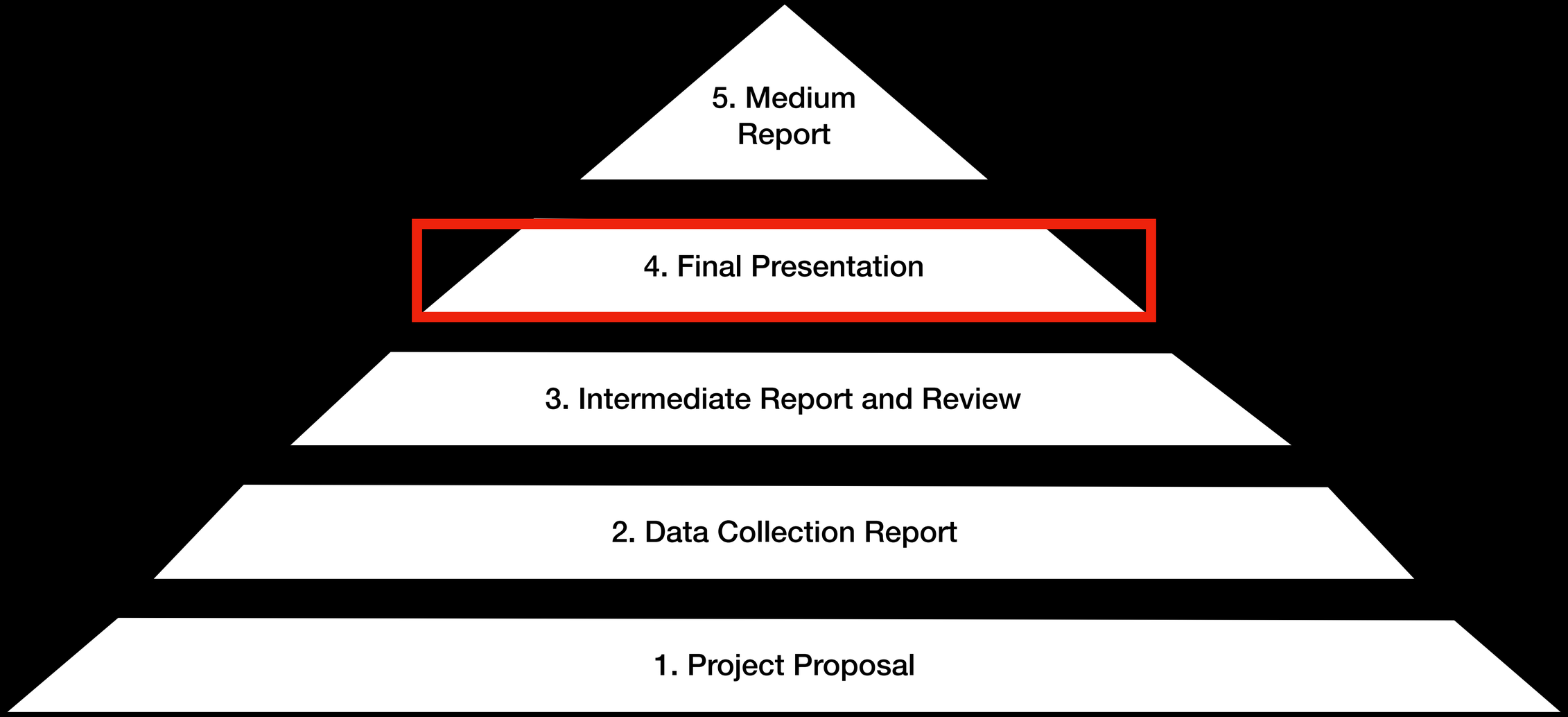
INST414 - Data Science Techniques

Medium Posts are **due no later than December 8**

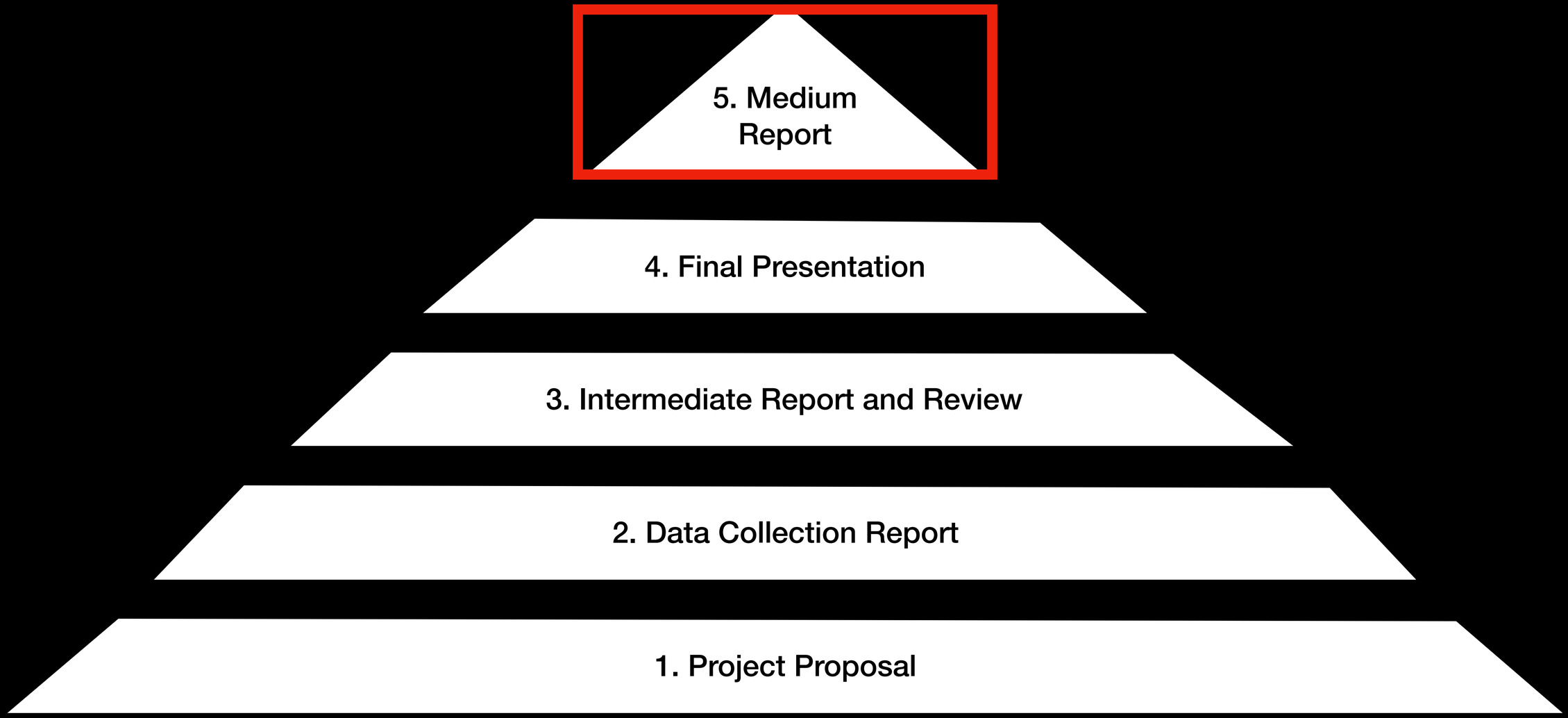
After this date, all unsubmitted assignments will get 0s



Semester Project



These are cancelled



We will discuss the final report next week

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Account
Dashboard
Courses
Calendar
Inbox
History
Commons
CourseExp
Help
EMT
Course Policies

Spring 2023
Home
Announcements
Assignments
Discussions
Grades
People
Pages
Files
Syllabus
Outcomes
Modules
BigBlueButton
Collaborations
Chat
Panopto Recordings
New Analytics
Clickers
Course Reserves
Adobe Creative Cloud
Quiz Extensions
Ally Course Accessibility Report

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Account
Dashboard
Courses
Calendar
Inbox
History
Commons
CourseExp
Help
EMT
Course Policies

Spring 2023
Home
Announcements
Assignments
Discussions
Grades
People
Pages
Files
Syllabus
Outcomes
Rubrics
BigBlueButton
Collaborations
Chat
Panopto Recordings
New Analytics
Clickers
Course Reserves
Adobe Creative Cloud
Quiz Extensions
Ally Course Accessibility Report

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Account
Dashboard
Courses
Calendar
Inbox
Portfolium
History
CourseExp
Help
EMT
Course Policies

Spring 2023
Home
Announcements
Assignments
Discussions
Grades
People
Pages
Files
Syllabus
Outcomes
Rubrics
Quizzes
Modules
BigBlueButton
Collaborations
New Analytics
Clickers
Course Reserves
Adobe Creative Cloud
Quiz Extensions
Ally Course Accessibility Report

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Account
Dashboard
Courses
Calendar
Inbox
Portfolium
History
Commons
Help
EMT
Course Policies

Spring 2023
Home
Announcements
Assignments
Discussions
Grades
People
Pages
Files
Syllabus
Outcomes
Rubrics
Quizzes
Modules
BigBlueButton
Collaborations
Chat
Course Reserves
Adobe Creative Cloud
Quiz Extensions
Ally Course Accessibility Report

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Semester Project - 5 - Final [Medium] Report

Published | Edit

Your project will culminate with a final report of at least **1,500 to 2,000 words**, with an additional **500 words** per additional student team member, posted to the class's Medium publication. This report should cover the breadth of your work and the major accomplishments of your project to a sufficient level of detail that someone should have a decent idea of how to replicate it. In particular, be sure to justify any major design decisions you made.

If you work in a group of more than 1 student, then you should include an appendix that makes it painfully clear what part of the project was worked on by each student. If the distribution of work was complicated, a table may be a good idea.

Note this appendix and any related work are exempt from the word limit identified above.

Content

Your final report must include the following:

1. Provide a succinct title for your project.
2. Explain the decision you are supporting, so the reader knows what your actionable insight is. If you prepared a thorough proposal and intermediate report, then you may be able to borrow some material from there.
3. Explain what data you explored. Where did it come from, how did you process it? If your data collection report was thorough, you can likely reuse much of this material.
4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after

Related Items

SpeedGrader™

How to use UMD Canvas

Textbooks

Adopt Textbook

4. What is the key ideas from the course did your project build on? This should describe the rational behind and what methods you applied.
5. Explain what analysis you did to support the actionable insight you wanted to extract. What models did you apply, how did you evaluate them, etc.?
6. Explain what answer you have for the stakeholder who needs to make the decision you are supporting. This is often greatly aided through charts of experiments, but you should also include a clear description of your insights.
7. Conclude with a discussion of the limitations of your project, the data, your analysis, ethics, etc.

Once complete, publish your Medium article, submit a link to the article, and submit it to the class publication.

Tag your story as "inst414spr22final" - Articles with this tag will automatically be added to the final-report tab after I accept on the class publication.

Source Code Submission via GitHub

Beyond the report describing your work, you must also submit the associated code you developed for this analysis. To submit this code, you should create a GitHub repository, and include a link to that repository in your final report on Medium.

If you do not want to share your code publicly, you have the option of creating a private GitHub repository and adding me as a collaborator, so I can see your code. My username on GitHub is cbuntain. If you take this option, you must include the following line in your report:

- "Source code for this project is available on request."

GitHub has free tiers of service, and you are entitled to additional free resources on that platform as a student.

GitHub also has significant documentation for creating repositories of your code, which you can find here: https://docs.github.com/en/get-started/quickstart/create-a-repo

Grading Rubric

The rubric for this report is as follows: Rubric.doc

Points 100

Submitting a website url

This Module's Learning Objectives

Evaluation in Supervised Learning

Describe why accuracy may be a poor metric for imbalanced data

Calculate precision and recall for a supervised learning model

Differentiate macro- and micro-averaging for multi-class evaluation

Use k-fold cross-validation to estimate average performance

This Module's Learning Objectives

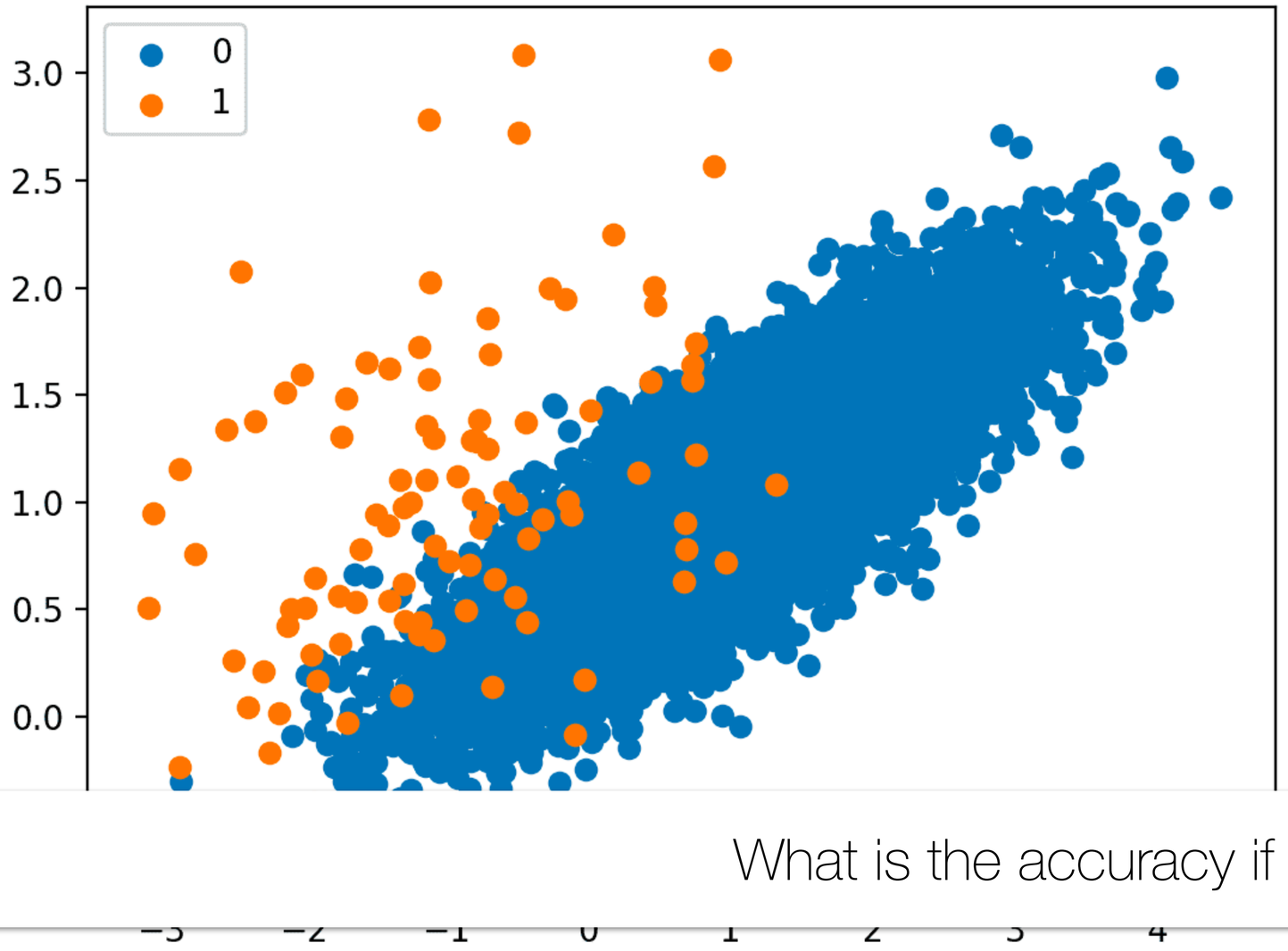
Evaluation in Supervised Learning

Describe why accuracy may be a poor metric for imbalanced data

Calculate precision and recall for a supervised learning model

Differentiate macro- and micro-averaging for multi-class evaluation

Use k-fold cross-validation to estimate average performance



What is the accuracy if we always predict 0?

This Module's Learning Objectives

Evaluation in Supervised Learning

Describe why accuracy may be a poor metric for imbalanced data

Calculate precision and recall for a supervised learning model

Differentiate macro- and micro-averaging for multi-class evaluation

Use k-fold cross-validation to estimate average performance

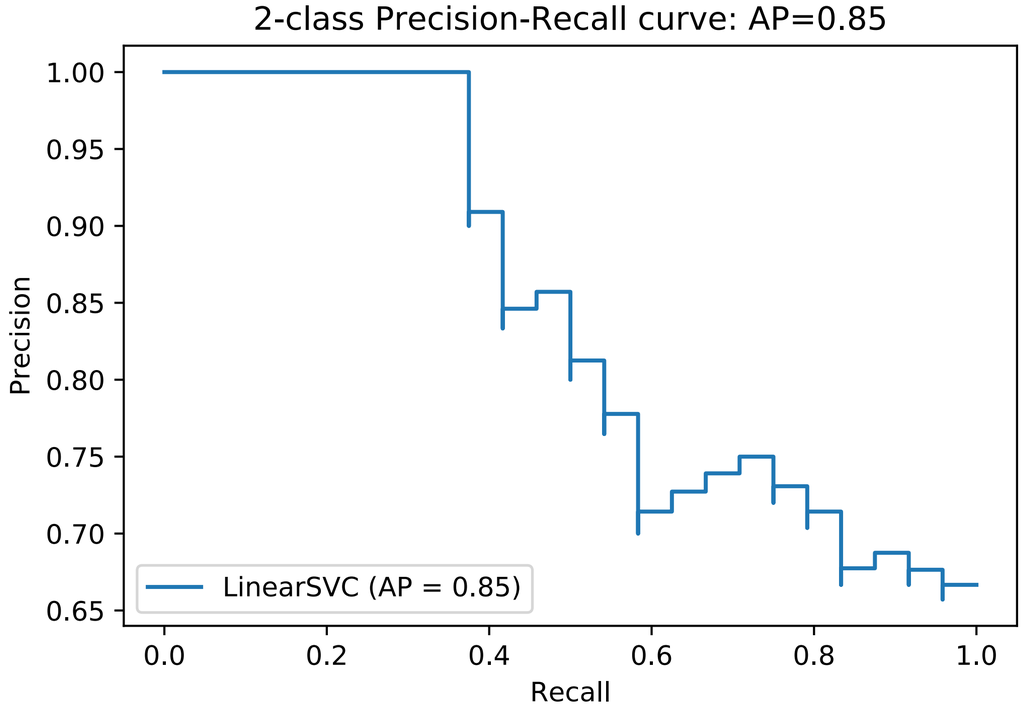
$$\text{Recall} = \frac{TP}{TP + FN}$$

What does *recall = 1.0* mean?

$$\text{Precision} = \frac{TP}{TP + FP}$$

What does *precision = 1.0* mean?

Precision-Recall Tradeoff



- Recall-oriented tasks:
  - Tumor detection
  - Legal discovery
  - (False Negatives are more costly.)
- Precision-oriented tasks:
  - Search engine ranking
  - Sentiment classification
  - Customer-facing tasks (users remember failures!)
  - (False Positives are more costly.)

Combining Precision and Recall with F-score

- $F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
- $F_\beta = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}}$
- β allows adjustment of the metric to control the emphasis on recall vs precision:
  - Precision-oriented: small β
  - Recall-oriented: large β

This Module's Learning Objectives

Evaluation in Supervised Learning

Describe why accuracy may be a poor metric for imbalanced data

Calculate precision and recall for a supervised learning model

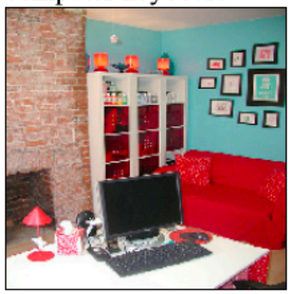
Differentiate macro- and micro-averaging for multi-class evaluation

Use k-fold cross-validation to estimate average performance

Old label: pier
ReaL: dock; pier; speedboat; sandbar; seashore

Old label: hammer
ReaL: screwdriver; hammer; power drill; carpenter's kit

Old label: monitor
ReaL: mouse; desk; desktop computer; lamp; studio couch; monitor; computer keyboard



Old label: zucchini
ReaL: broccoli; zucchini; cucumber; orange; lemon; banana

Old label: ant
ReaL: ant; ladybug

Old label: quill
ReaL: feather boa

Old label: water jug
ReaL: water bottle

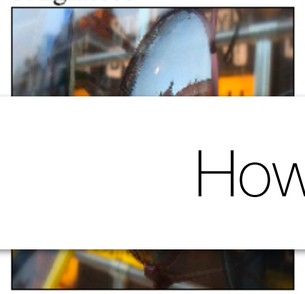
Old label: chain
ReaL: necklace

Old label: purse
ReaL: wallet

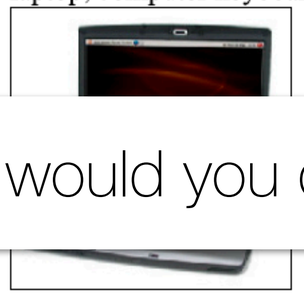
Old label: passenger car
ReaL: school bus

Old label: sunglass
ReaL: sunglass; sunglasses

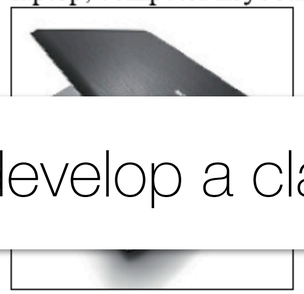
Old label: sunglasses
ReaL: sunglass; sunglasses



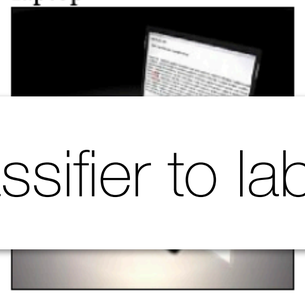
Old label: laptop
ReaL: notebook; laptop; computer keyboard



Old label: notebook
ReaL: notebook; laptop; computer keyboard



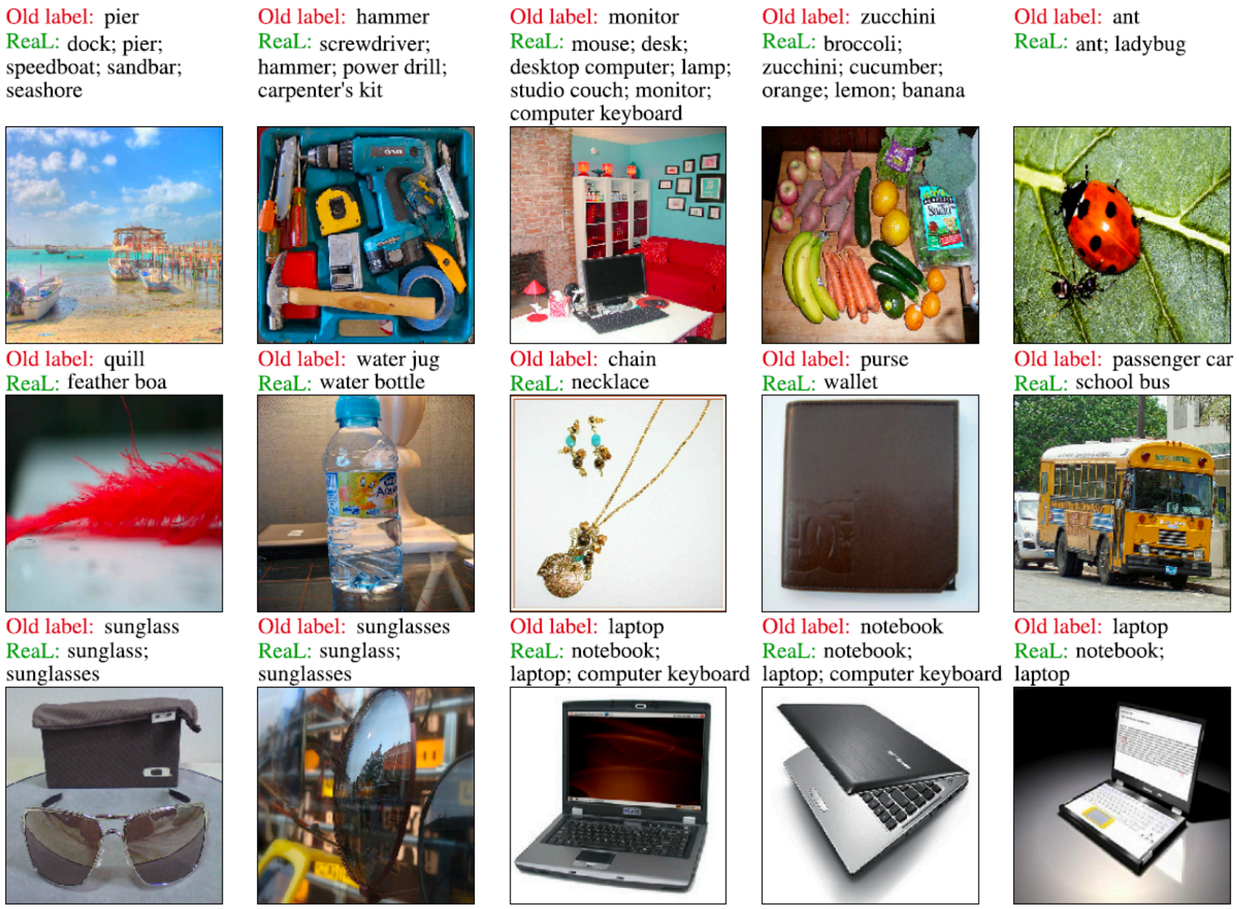
Old label: laptop
ReaL: notebook; laptop



How would you develop a classifier to label these images?

Multi-Class Classification

- An extension to binary classification
  - How are binary and multi-class different?
- Multiple ways to solve this problem:
  - Train a single classifier that supports different outcomes (e.g., decision tree)
  - Train many one-vs-rest classifiers
  - Train many more one-vs-one classifiers (for all pairs of labels)



Is this a pier (y=1) or not (y=0)?

Is this a drill (y=1) or not (y=0)?

One vs. rest strategy

Is this a monitor (y=1) or not (y=0)?

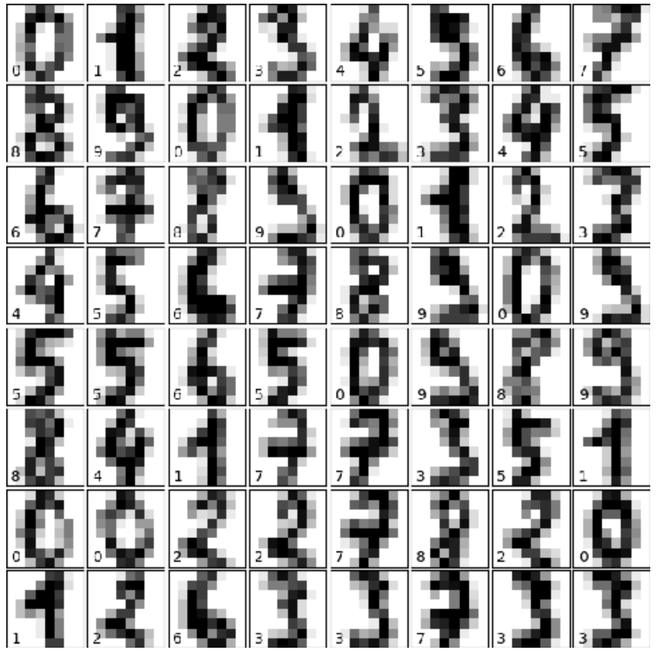
Is this a banana (y=1) or not (y=0)?

Is this a bus (y=1) or not (y=0)?

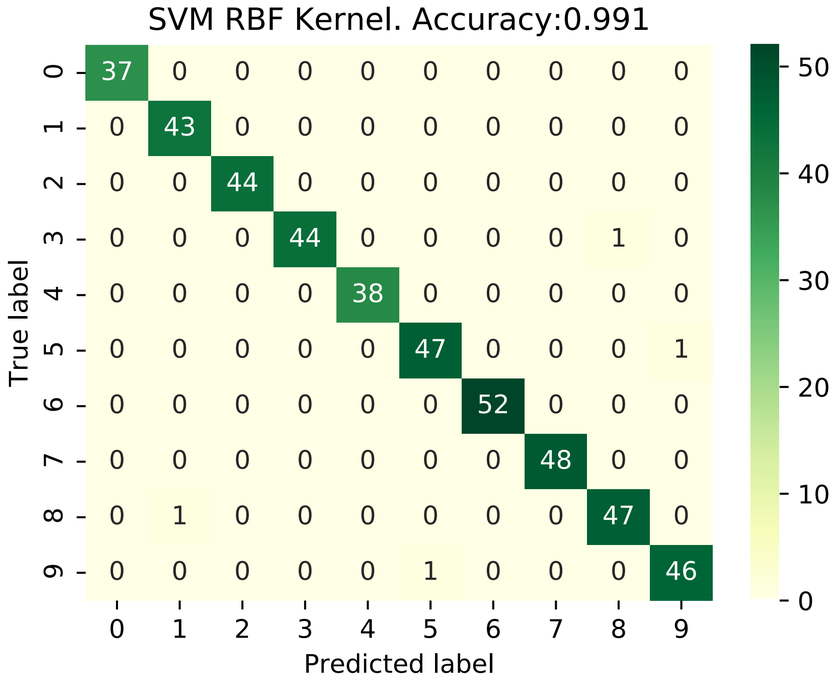
Evaluation in Multi-Class Classification

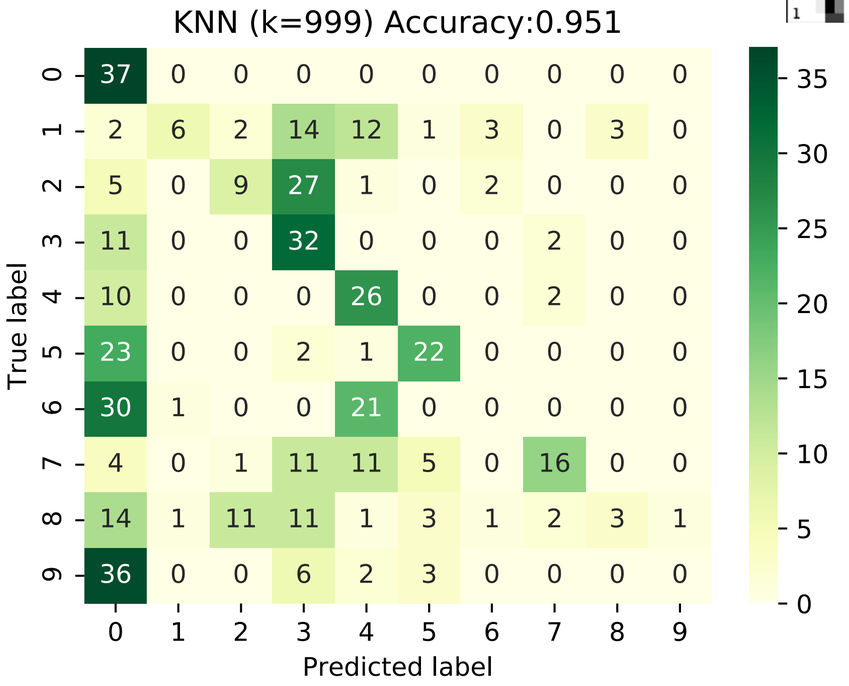
- Multi-class evaluation is an extension of the binary case
  - A collection of true vs. predicted binary outcomes, one per class
  - Confusion matrices are especially useful

Multi-Class Confusion Matrix



MNIST Dataset

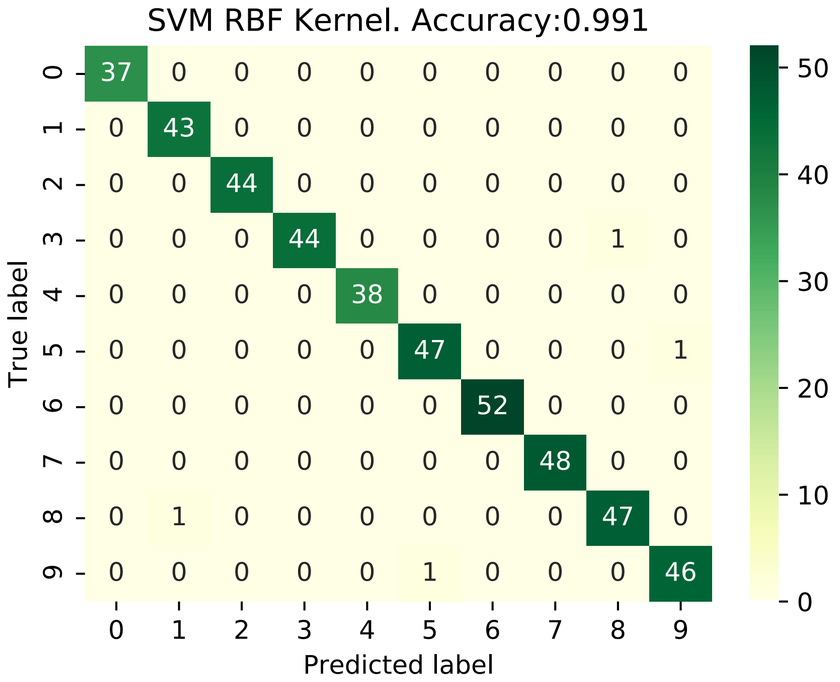




Multi-Class Classification

- Overall evaluation metrics are average across classes
  - But there are different ways to average

Average across this matrix



Micro-Average

- Each element has equal weights
  - Aggregate outcomes across all classes (count correct labels in each class)
  - Compute metric with aggregated outcomes (count correct / total count)

Class	Predicted Class	Correct?
Orange	Lemon	0
Orange	Lemon	0
Orange	Apple	0
Orange	Orange	1
Orange	Apple	0
Lemon	Lemon	1
Lemon	Apple	0
Apple	Apple	1
Apple	Apple	1

Micro Average Recall:
4 / 9 = 0.44

Largest classes have the most influence!

jupyter prediction_model_01.c /2023

Logout

File Edit View Language current mode

```
imdb_id,predicted,actual genres,correct?
tt10880158,Sci-Fi,['Comedy'],0
tt0462573,Documentary,"['Drama', 'Romance']",0
tt2160413,Adventure,['Drama'],0
tt6032704,Action,['Horror'],0
tt2198235,Documentary,"['Action', 'Crime', 'Drama']",0
tt0319829,History,"['Comedy', 'Drama', 'Music']",0
tt0206917,Reality-TV,"['Drama', 'Music', 'Romance']",0
tt6048930,Animation,"['Comedy', 'Fantasy', 'Romance']",0
tt4913966,Music,"['Horror', 'Mystery', 'Thriller']",0
tt7090140,Biography,"['Drama', 'Music']",0
tt12709774,Comedy,['Action'],0
tt1885299,Short,"['Comedy', 'Drama']",0
tt12458678,Music,"['Comedy', 'Horror']",0
tt3416742,News,"['Comedy', 'Horror']",0
tt4842646,Animation,"['Biography', 'Drama', 'History']",0
tt1374990,Reality-TV,"['Action', 'Animation', 'Drama']",0
tt0762114,Comedy,"['Comedy', 'Romance']",1
tt0251075,Western,"['Comedy', 'Sci-Fi']",0
tt15053518,War,"['Drama', 'Horror', 'Music']",0
tt14010416,Crime,['Action'],0
tt5300736,Crime,['Horror'],0
tt0396171,Musical,"['Crime', 'Drama', 'Fantasy']",0
tt4591226,Horror,['Horror'],1
tt0881913,Family,"['Drama', 'Family']",1
tt5741008,History,['Comedy'],0
tt1837569,Sport,"['Drama', 'Thriller']",0
tt1188113,Short,"['Drama', 'Thriller']",0
tt11197858,Biography,['Documentary'],0
tt1462901,Crime,"['Comedy', 'Drama']",0
tt10840042,Animation,"['Drama', 'Thriller']",0
tt10350922,Family,"['Action', 'Comedy', 'Horror']",0
tt0760329,Comedy,"['Adventure', 'Family', 'Fantasy']",0
tt0309593,Biography,"['Horror', 'Thriller']",0
tt0369436,Short,"['Comedy', 'Drama', 'Romance']",0
tt3344694,Comedy,"['Drama', 'Romance']",0
tt8589474,Adventure,"['Drama', 'Thriller']",0
tt1684558,Adventure,"['Comedy', 'Drama', 'Romance']",0
tt6405208,Romance,"['Action', 'Drama', 'Thriller']",0
tt0304912,Comedy,['Comedy'],1
tt0156376,Romance,"['Comedy', 'Romance']",1
tt0420740,Short,"['Drama', 'Thriller']",0
tt1105740,Western,['Drama'],0
tt3678656,Mystery,"['Comedy', 'Drama']",0
```

Micro-averaged accuracy: Sum“correct?” column / rows in the dataset

Why might this micro-averaging be problematic?

What happens when the classes are imbalanced?

Macro-Average

- Each class has equal weights
  - First compute metric within each class
  - Then average resulting metrics across classes

Class	Predicted Class	Correct?
Orange	Lemon	0
Orange	Lemon	0
Orange	Apple	0
Orange	Orange	1
Orange	Apple	0
Lemon	Lemon	1
Lemon	Apple	0
Apple	Apple	1
Apple	Apple	1

Class	Recall
Orange	1/5 = 0.2
Apple	2/2 = 1.0
Lemon	1/2 = 0.5

Macro Average Recall:
(0.2 + 0.5 + 1.0) / 3 = 0.57

All classes have the same influence!

Imbalance and Macro- vs. Micro-Averaging

- If each class has about the same number of instances, what happens?
  - micro-average < macro-average?
  - micro-average > macro-average?
  - micro-average ≈ macro-average?
- Why?

Imbalance and Macro- vs. Micro-Averaging

- If your data is imbalanced, and…
- You want to weight your metric toward the largest ones
  - Use micro-averaging
  - Where might you want this weighting?
- You want to weight your metric toward the smallest ones
  - Use macro-averaging
  - Where might you want this weighting?

Imbalance and Macro- vs. Micro-Averaging

- Rules of Thumb:
  - If micro-average << macro-average, then examine the larger classes.
  - If micro-average >> macro-average, then examine the smaller classes.

This Module's Learning Objectives

Evaluation in Supervised Learning

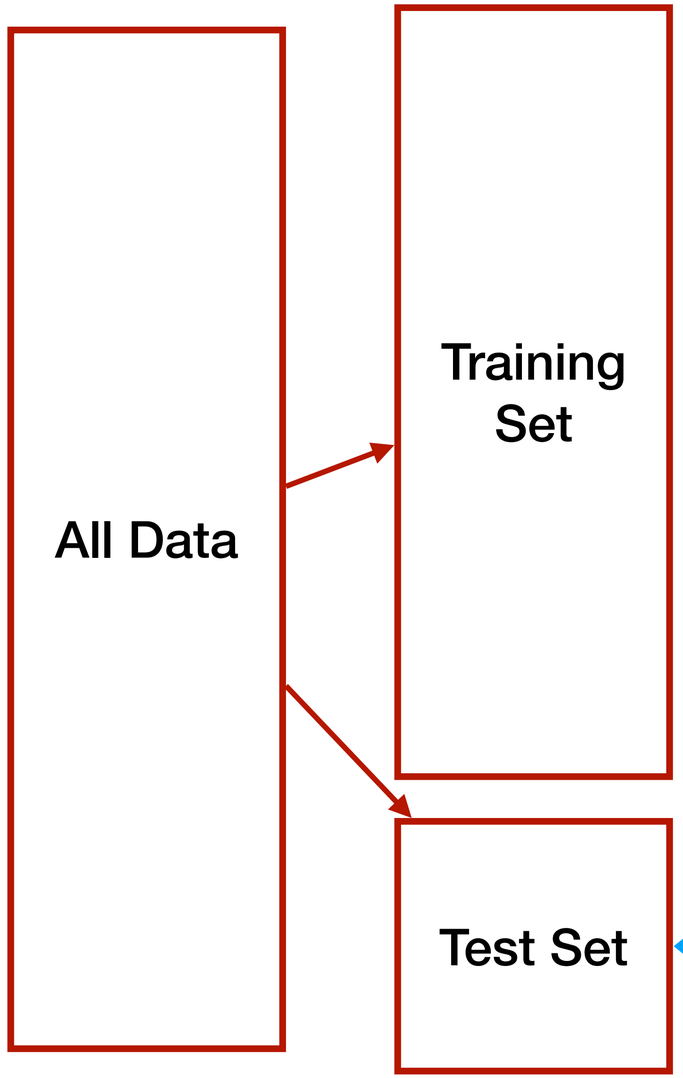
Describe why accuracy may be a poor metric for imbalanced data

Calculate precision and recall for a supervised learning model

Differentiate macro- and micro-averaging for multi-class evaluation

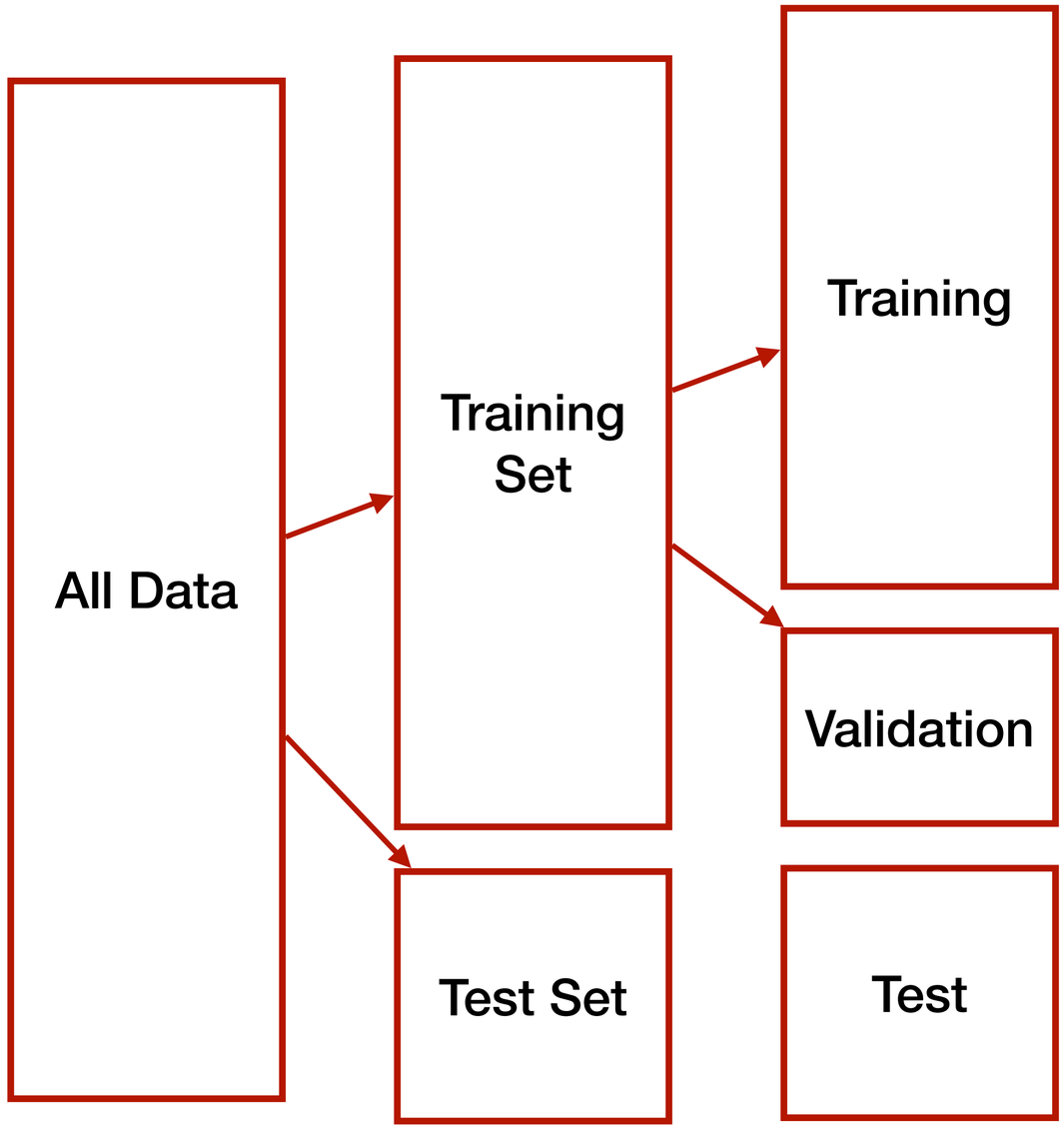
Use k-fold cross-validation to estimate average performance

Model Selection: Training-Test Split



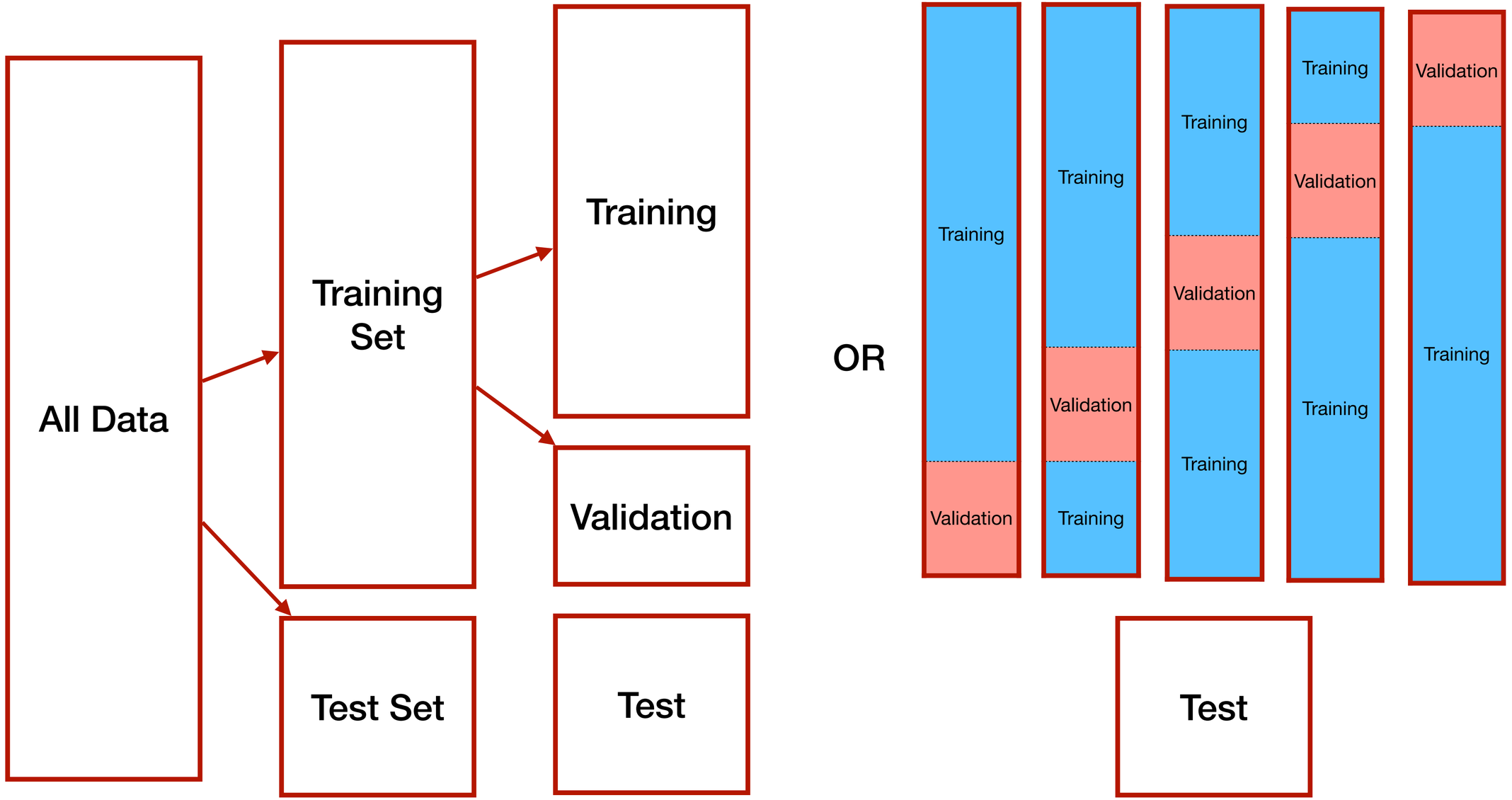
Now we lose this data from our training set… 👎

Model Selection: Training-Validation-Test Split

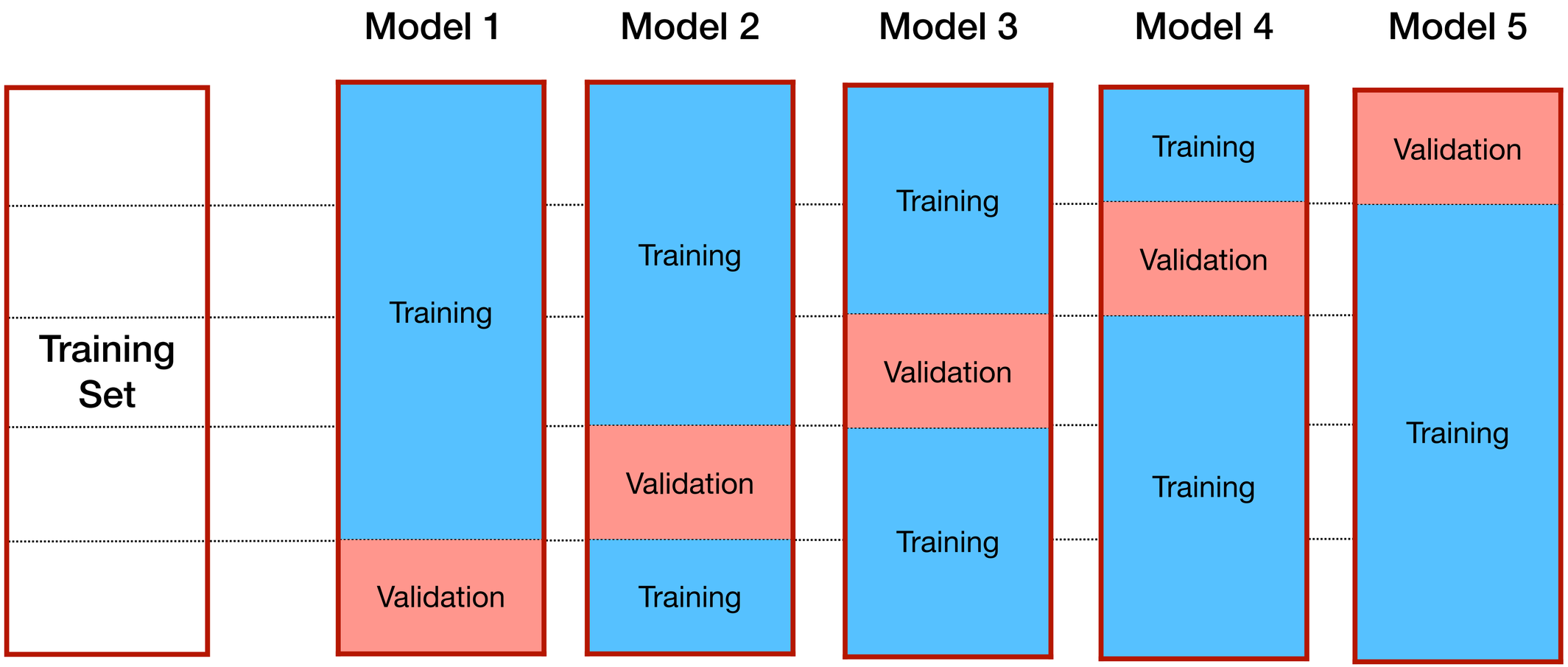


Now we lose this data from our training set… 👎

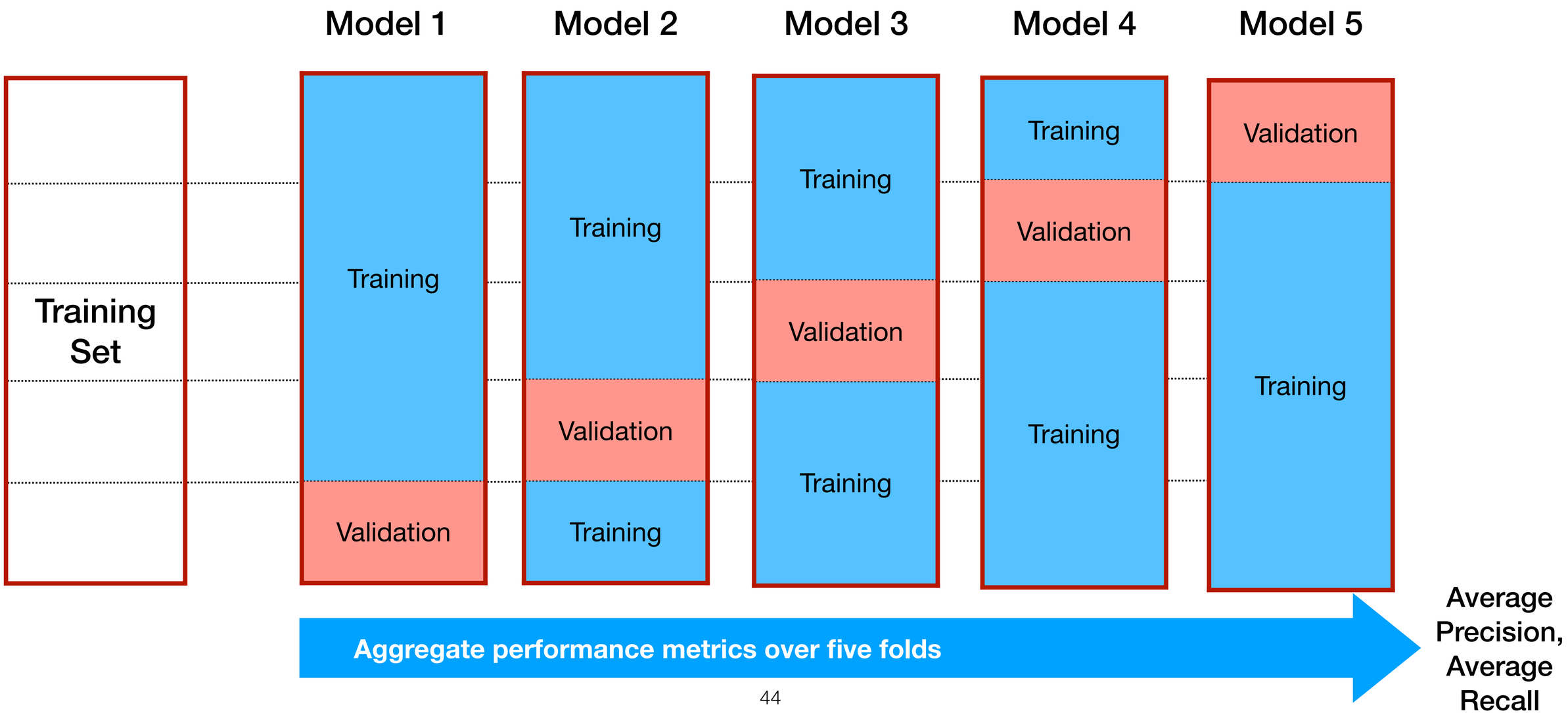
Model Selection: Training-Validation-Test Split

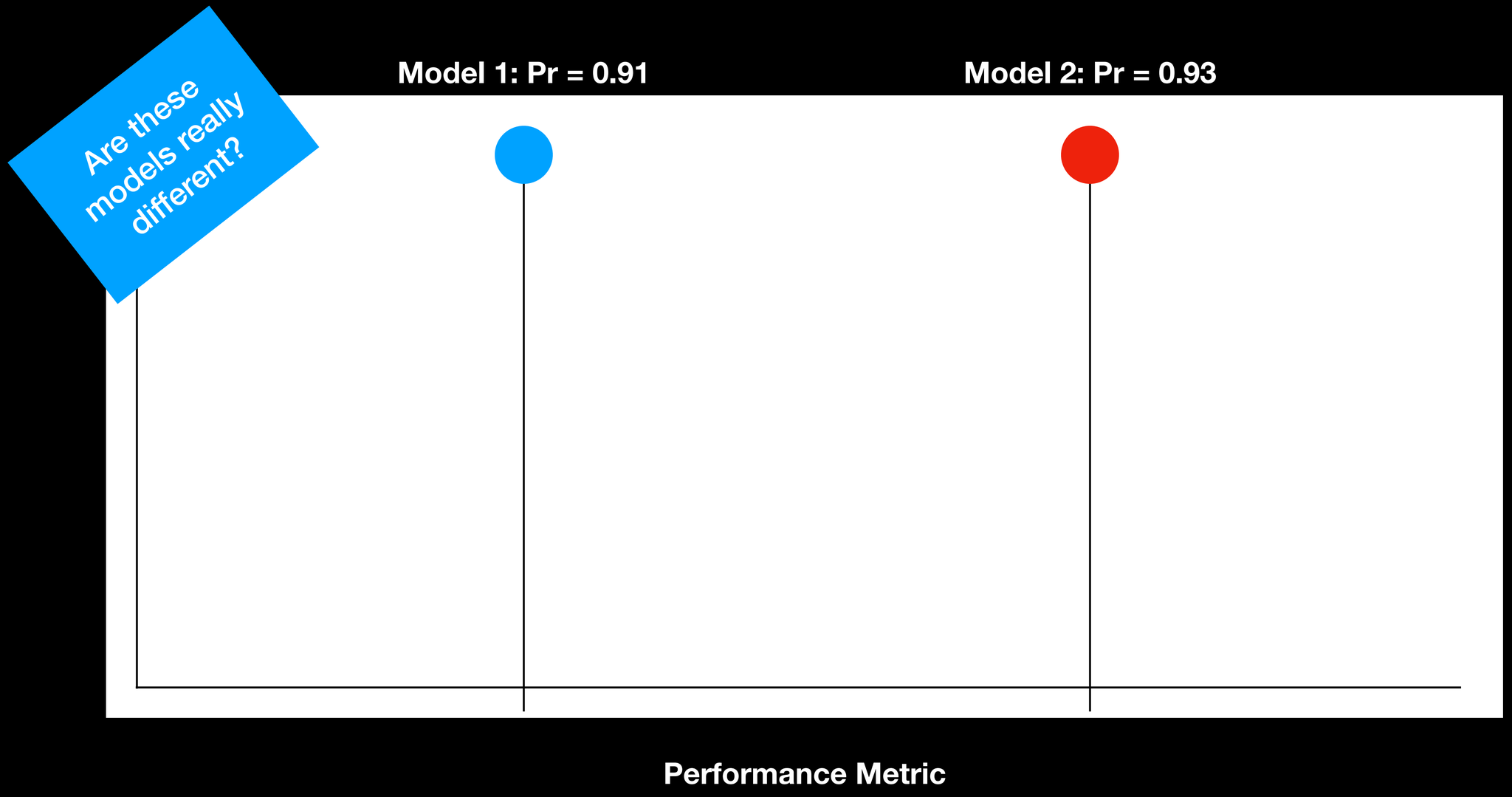


Five-fold Cross Validation

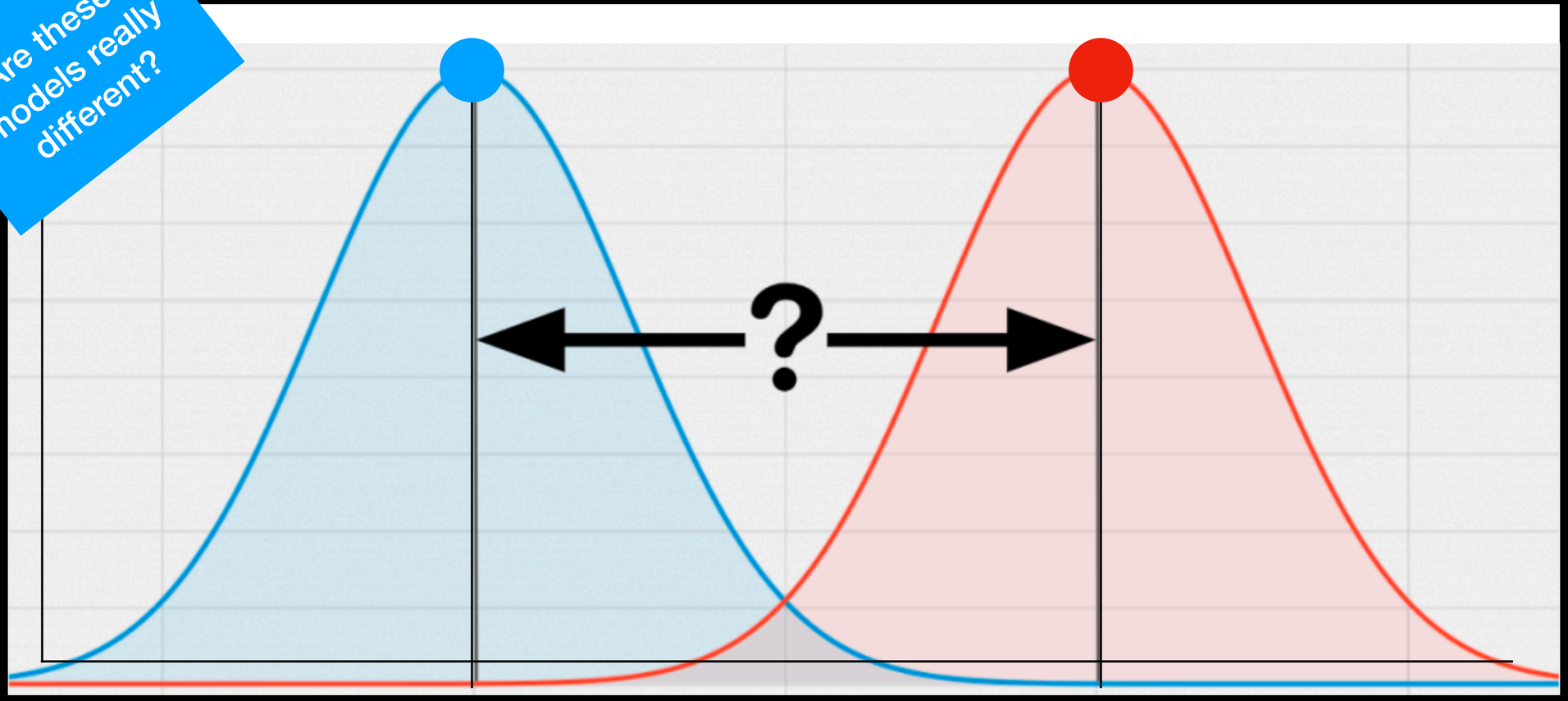


Five-fold Cross Validation

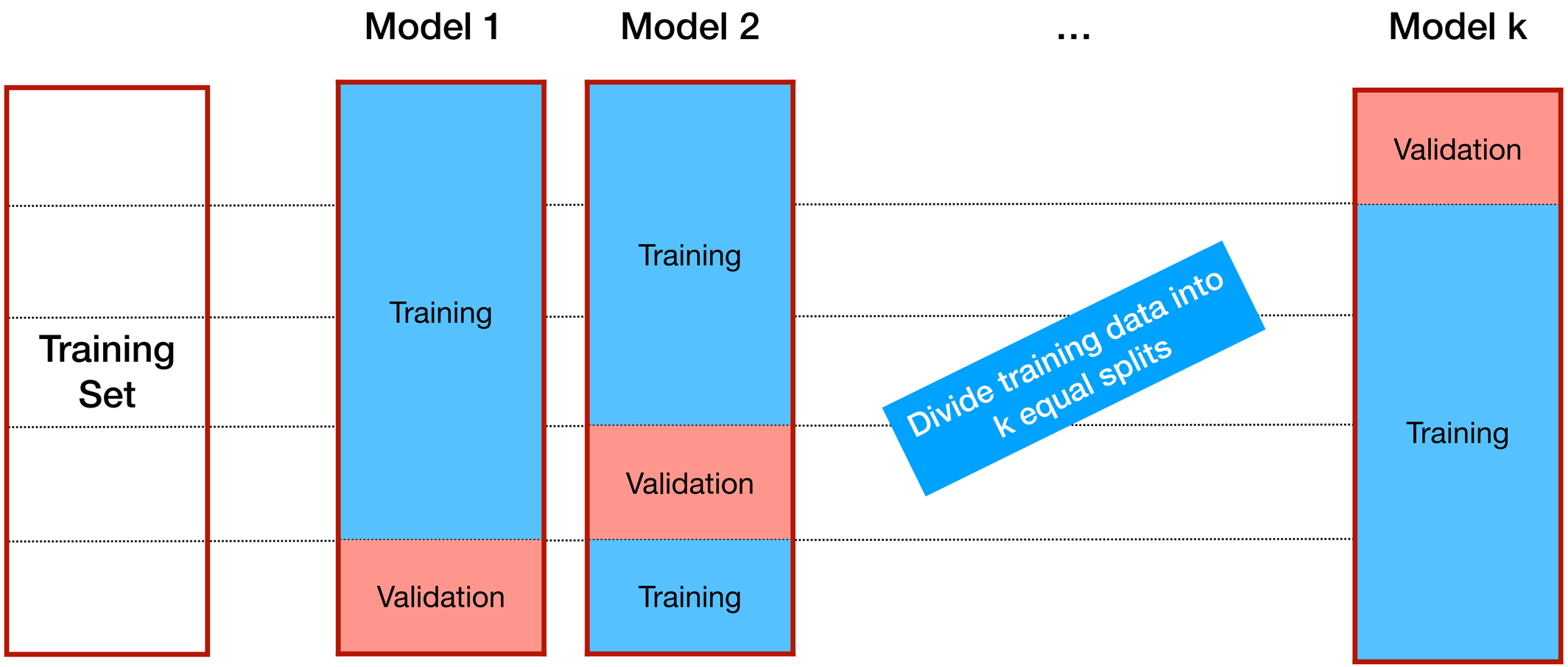



Are these models really different?
Model 1: Pr = 0.91
Model 2: Pr = 0.93
Performance Metric

Are these models really different?



k-fold Cross Validation



Concluding Notes on Model Evaluation and Selection

- Accuracy is often not the right evaluation metric for many real-world data science tasks
  - False positives and false negatives may need to be treated very differently
  - Make sure you understand the needs of your application and choose an evaluation metric that matches your application, user, or business goals

Concluding Notes on Model Evaluation and Selection

- Examples of additional evaluation methods include:
  - Learning curve: how much does accuracy (or other metric) change as a function of the amount of training data?
  - Sensitivity analysis: How much does accuracy (or other metric) change as a function of key learning parameter values?

What questions do you have?

Prof. Cody Buntain | @codybuntain | cbuntain@umd.edu
Director, Information Ecosystems Lab