Audio & Speech
Week4: Baby Sounds Challenge

Prof. Chi-Chun Lee, Yi-Wen Liu

TAs:邱信豪、許暐彤、陳舫慶、陳靖杰

EE3662: Digital Signal Processing Lab

#Lab12 – Dec. 20, 2021

BIIC

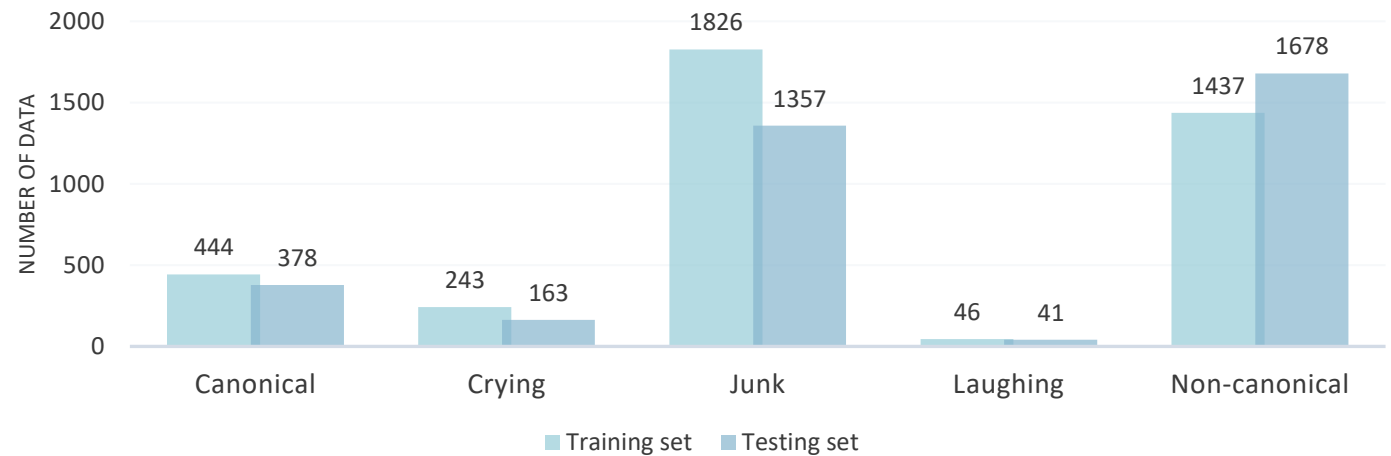# Dataset Description

- **Baby Sounds Challenge**
  - The INTERSPEECH 2019 Computational Paralinguistics Challenge

- **Datasets**
  - Training set: 3996 files
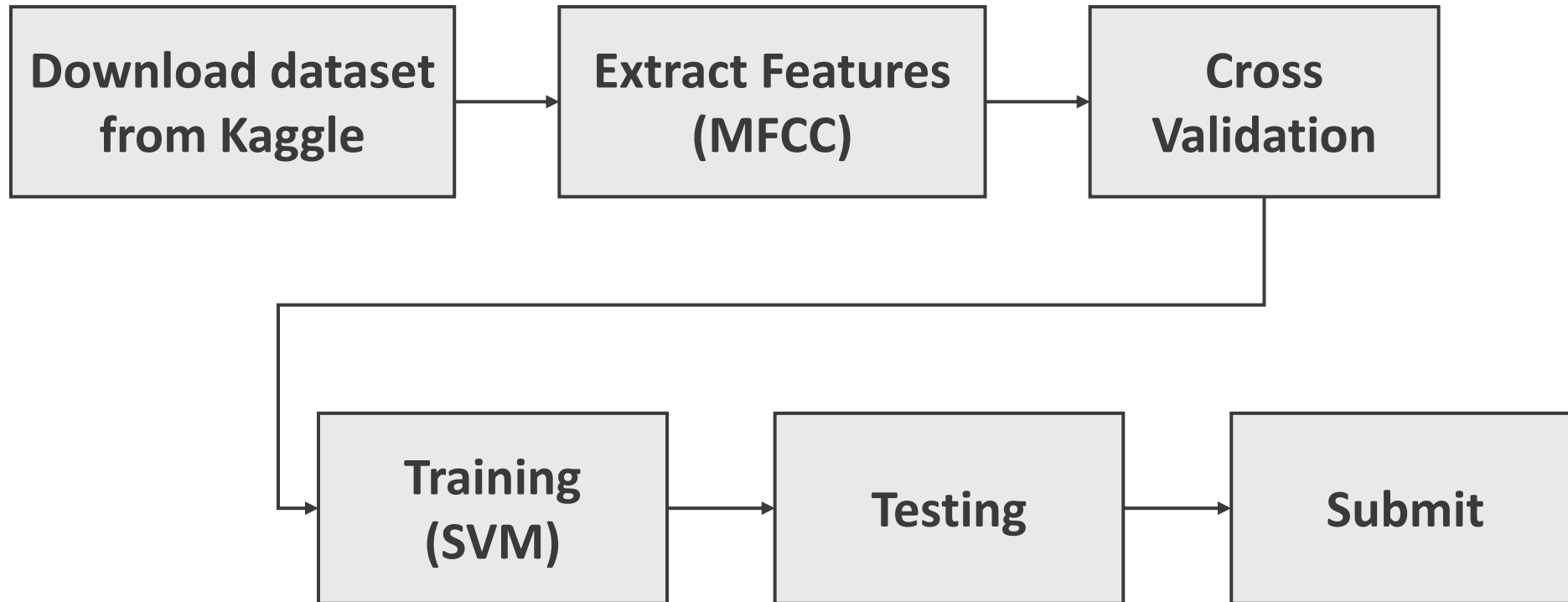 - Testing set: 3617 files

- **Download Datasets**
  - Kaggle → Data

Flow Chart

Kaggle

- Join this link and create your own account:
  - Create your **Team** name: using your student ID
  - One member one team!
  - We will give you competition link on 12/20.

Kaggle (Download Datasets)

- Click **Data**

Files

- **wav_train** - the training audio set, include 3996 samples.
- **wav_dev** - the test audio set, include 3617 samples.
- **label_raw_train.csv** - the training labels.
- **upload_sample.csv** - a sample csv file that you need to follow the format to submit.

```
kaggle competitions download -c dsplab-detecting-baby-sounds-2021
```

Data Explorer

87.42 MB

- ▸ 📁 Baby_Data
- ▥ label_raw_train.csv
- ▥ upload_sample.csv

< **label_raw_train.csv** (99.76 kB)

Detail   Compact   Column   2 of 2 columns ⌄

About this file

This file does not have a description yet.

A file_name		A label	
3996 unique values		Junk	46%
		Non-canonical	36%
		Other (733)	18%
train_0001.wav		Non-canonical	

Kaggle (Upload results)

- Click *submit predictions*
- Upload your results.csv
  - Results format:

	A	B
1	file_name	Predicted
2	devel_0001.wav	Crying
3	devel_0002.wav	Crying
4	devel_0003.wav	Crying
5	devel_0004.wav	Crying
6	devel_0005.wav	Crying
7	devel_0006.wav	Crying
8	devel_0007.wav	Crying
9	devel_0008.wav	Crying
10	devel_0009.wav	Crying

You can refer to upload_sample.csv

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

Message

Make a submission for **DSPlabTA**

You have 19 submissions remaining today. This resets 14 hours from now (00:00 UTC).
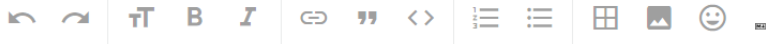
Step 1
Upload submission file

File Format
Your submission should be in CSV format. You can upload this in a zip/gz/rar/7z archive, if you prefer.

Number of Predictions
We expect the solution file to have 3617 prediction rows. This file should have a header row. Please see sample submission file on the data page.

Step 2
Describe submission

Briefly describe your submission
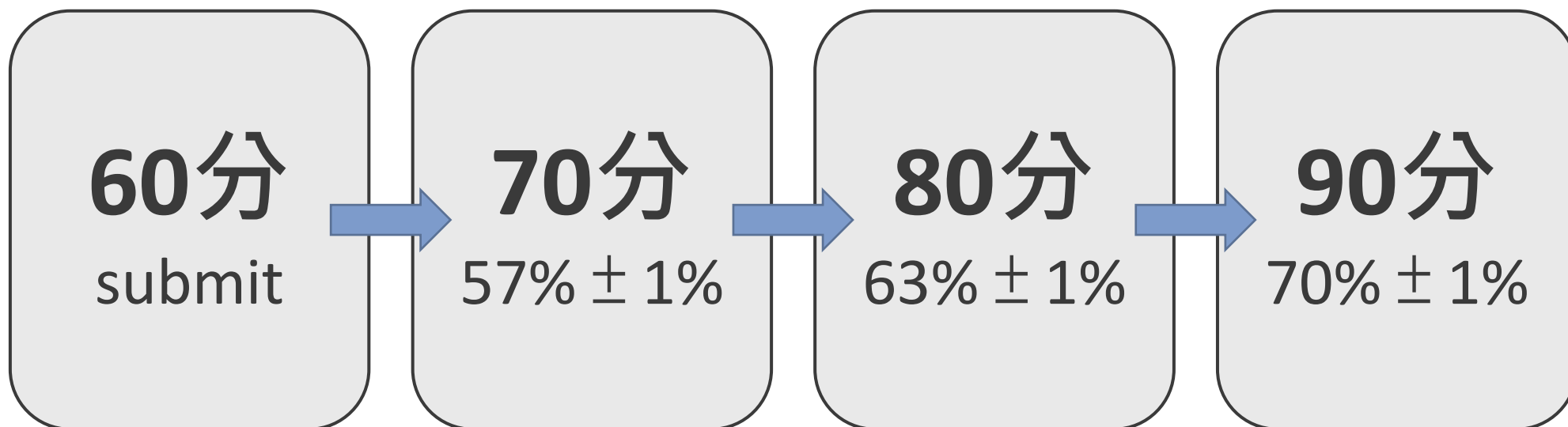
BIIC

Lab12 Grading

Kaggle Challenge Results (40%) + Report (60%)

Kaggle Challenge Grading

60分
submit

➤

70分
57%±1%

➤

80分
63%±1%

➤

90分
70%±1%

• **90分以上:** For every 1% improvement, you will get an extra 2 points

Challenge Hints

- Please use MFCC implemented in Lab9 by yourself to extract MFCC features.
- Kaggle has maximum daily submissions (5 times one day), so you can use SVM and Cross Validation in Lab11 to verify the training data
- Try different parameters (kernel, c, degree) to tune your model
- Try different features, statistics functions or classifiers to achieve higher performance
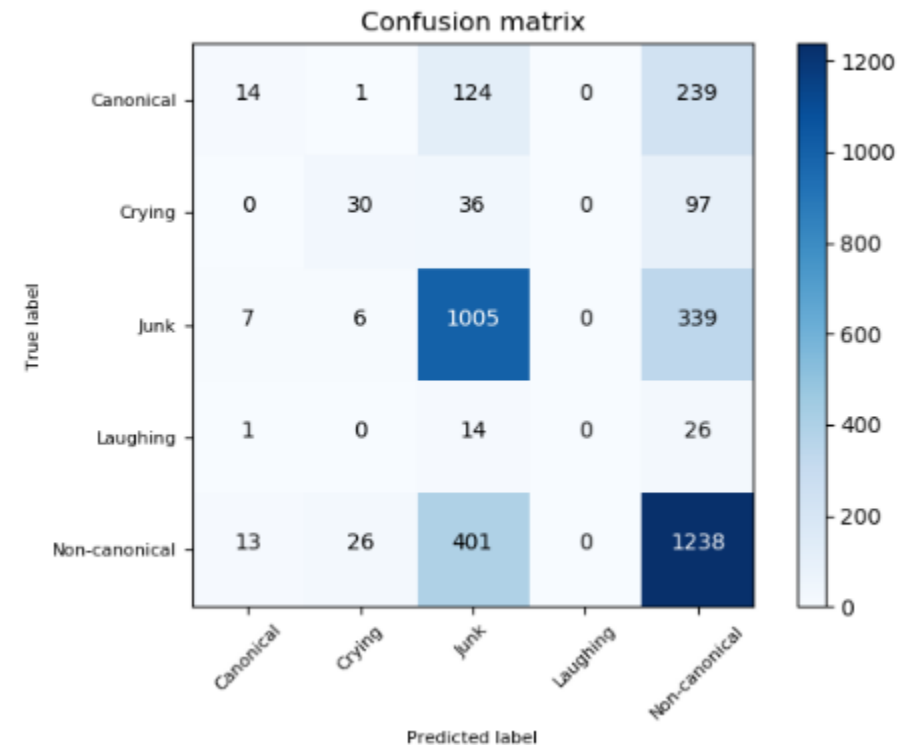
Report

- Try to explain your accuracies from your results table
- Using your best results to plot the confusion matrix (lab11)

Baby Sounds (%)						
	Canonical	Crying	Junk	Laughing	Non-canonical	UAR
SVM (rbf)	3.7	18.4	74.06	0	73.78	63.23
SVM (linear)						
SVM (poly)						
…						
…						

Confusion matrix

Report

- (Bonus) Analyses
    1. Try to explain why some categories perform worse? The data distribution seems quite imbalanced, try to solve the problems and compare the experimental results.
    2. What other problems did you meet? What is your solution? Any interesting findings?