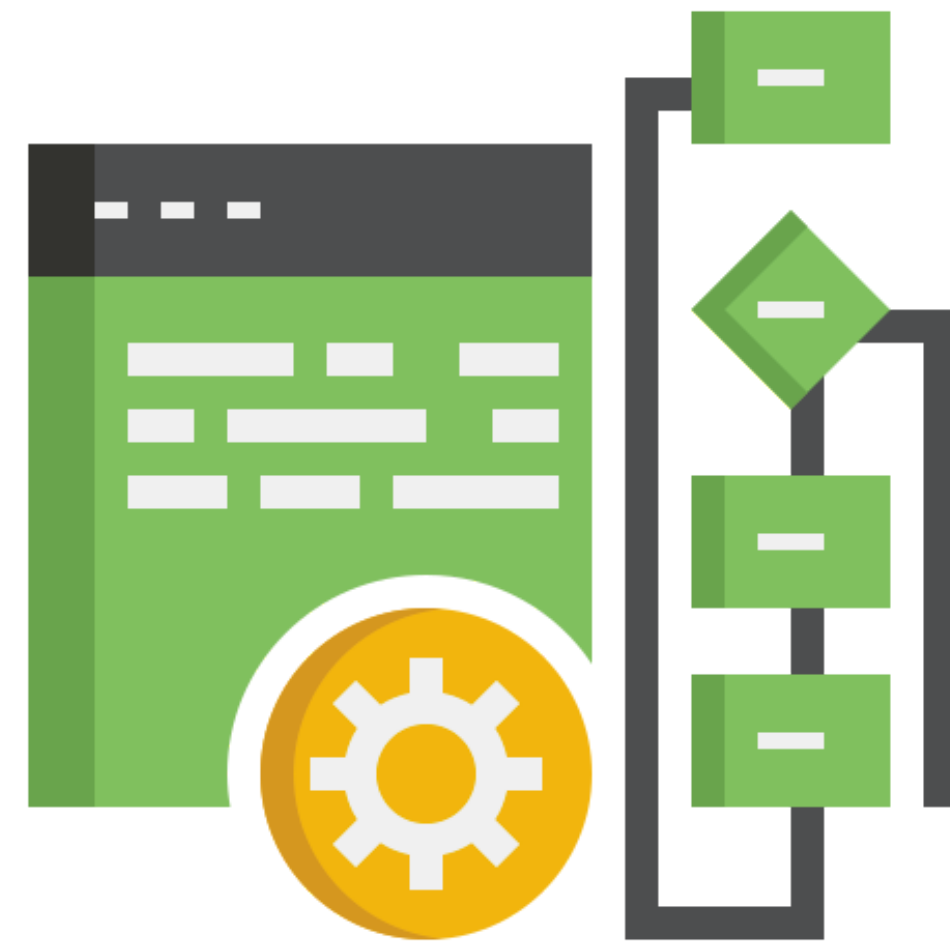


# Multi-class boosting with adversarial multi-arm bandits on incomplete views

Andrea Simeon  
(Mihajlović)

Sanja Brdar, Miloš Radovanović, Tatjana Lončar-Turukalo, Michelangelo  
Ceci and Gianvito Pio





## STSM

Short-term scientific mission

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.

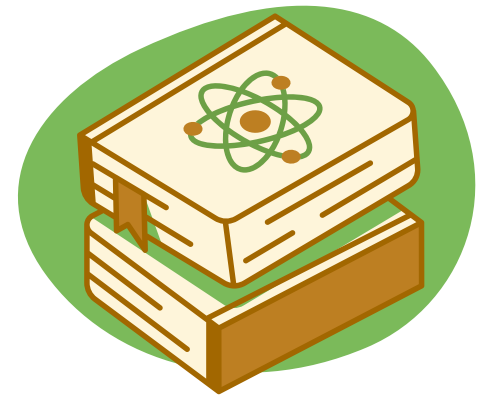


<https://www.cost.eu>  
<https://www.ml4microbiome.eu>



Funded by  
the European Union

# TOPIC OUTLINE



**MOTIVATION**



**ALGORITHM**

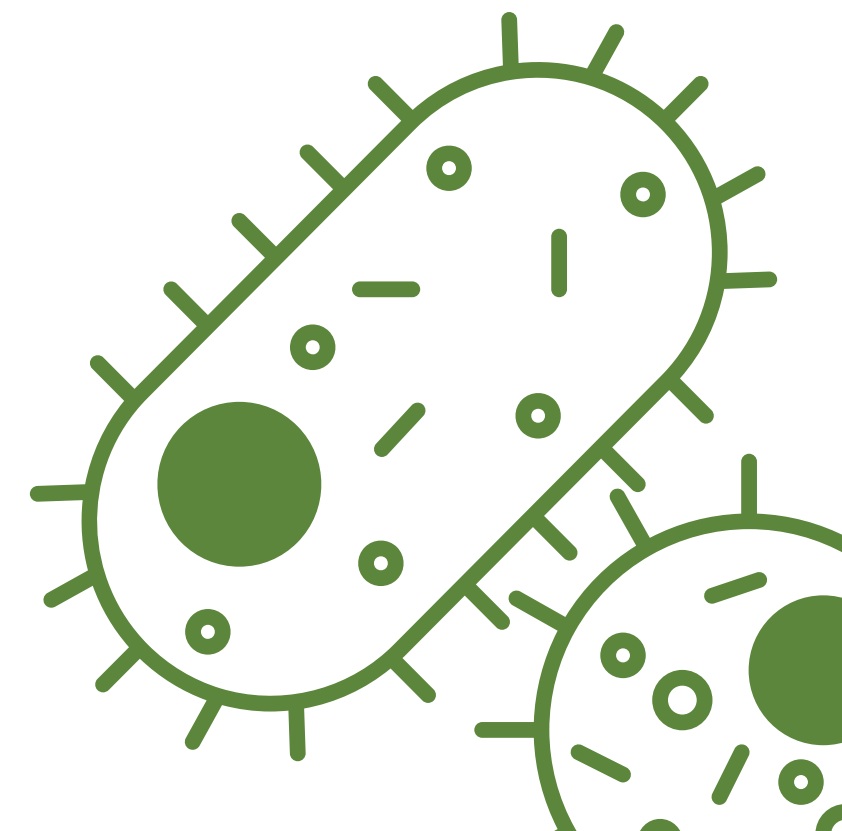
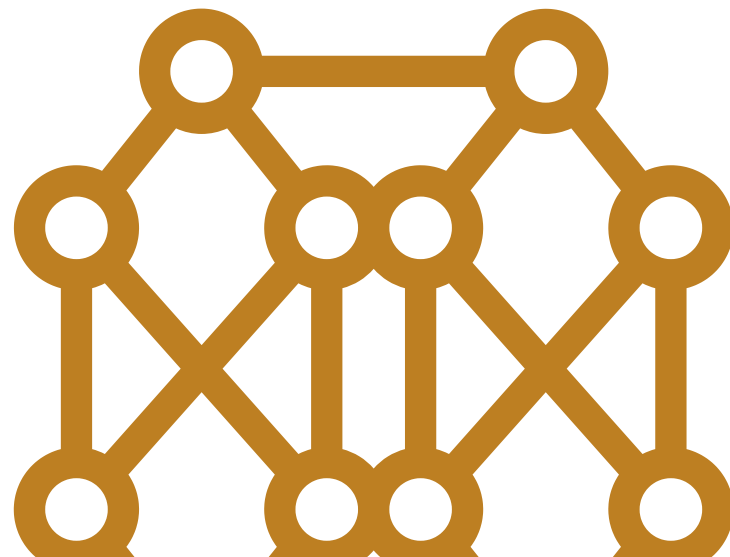
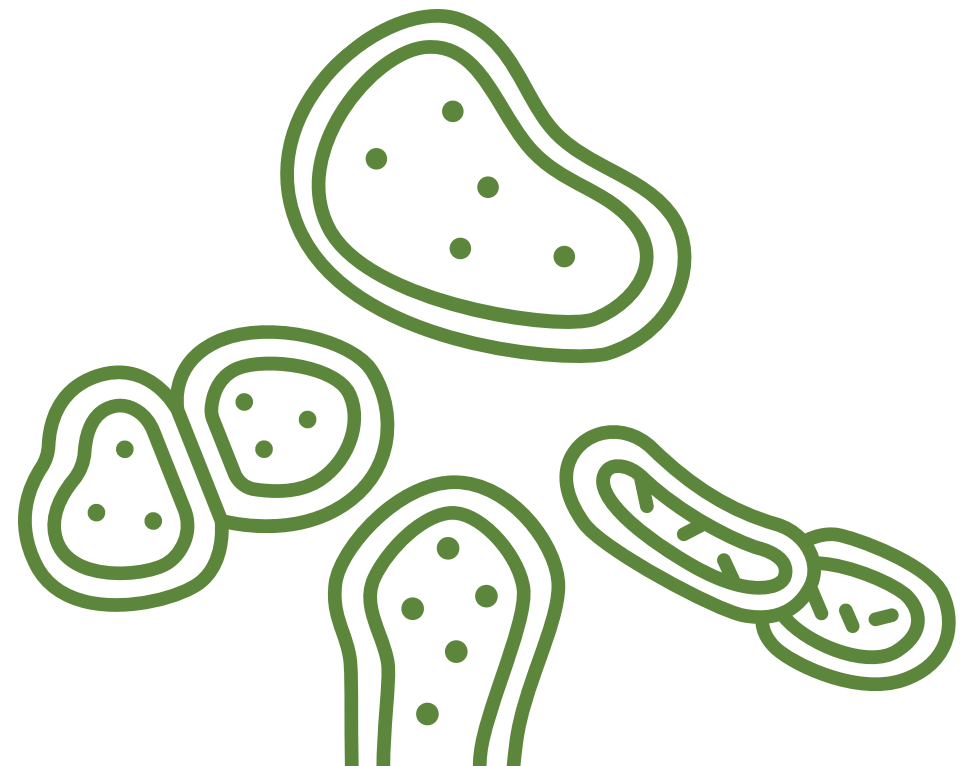


**RESULTS**



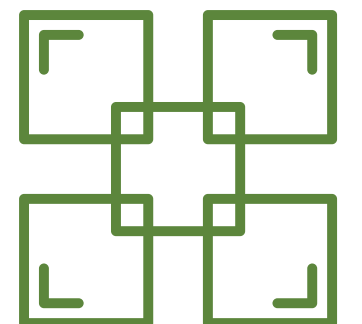
**CONCLUSION**

# MOTIVATION



# MICROBIOME

- Sequence data analysis
- Choose:
  - a. Sequencing technique: amplicon vs shotgun
  - b. Preprocessing technique: different pipelines
- Associate with diseases and conditions (ML classification)
- Dan et al. (2020) doi: [10.1080/19490976.2020.1747329](https://doi.org/10.1080/19490976.2020.1747329)
- Multi-view learning (fusion)
- Peng et al. (2018) doi: [10.1109/TNNLS.2016.2637881](https://doi.org/10.1109/TNNLS.2016.2637881)



# PROBLEM STATEMENT



## PROBLEM 1: MULTIPLE INCOMPLETE VIEWS

- Multiple measurements on the same data
- Incomplete in terms of samples across views

## PROBLEM 2: MULTI-CLASS

- More than two classes



# EXAMPLE

## Complete views

	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$
$s_1$	5	10	2
$s_2$	1	1	2
$s_3$	7	4	1
$s_4$	5	4	3

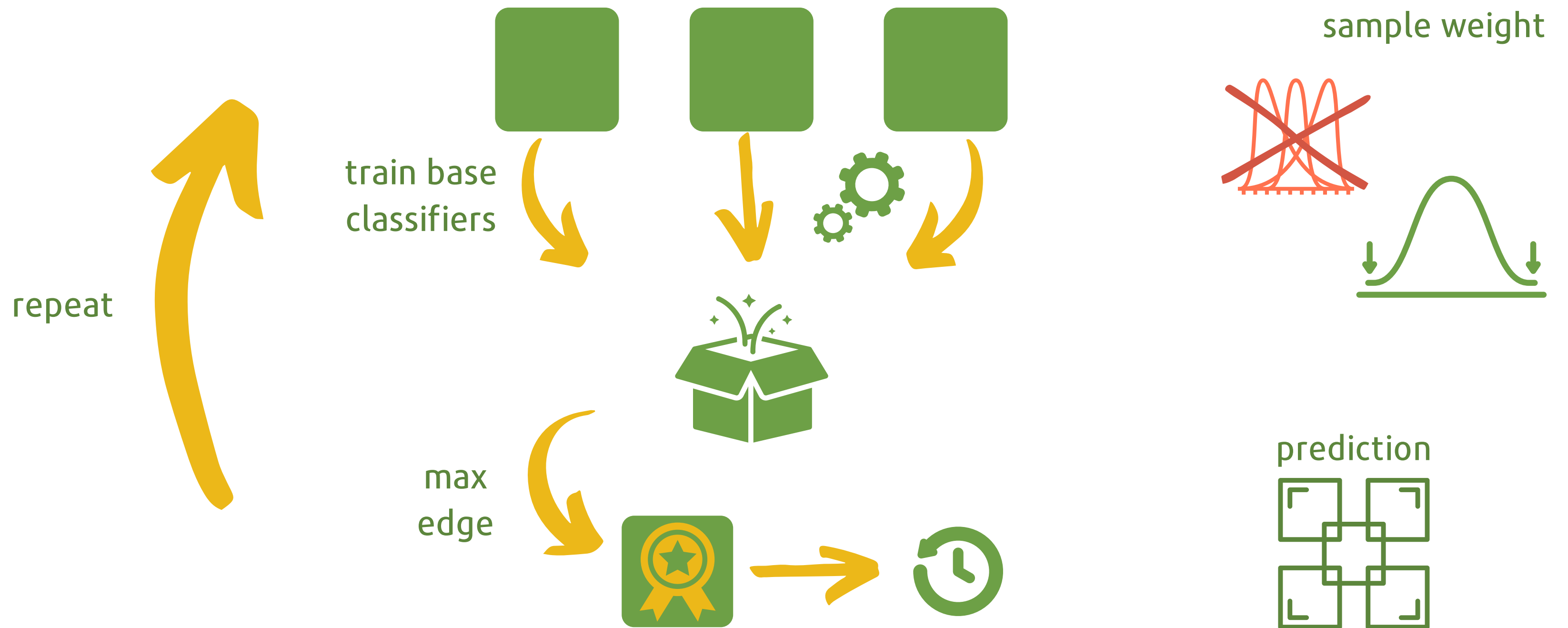
	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$
$s_1$	10	7	2
$s_2$	1	2	2
$s_3$	3	4	7
$s_4$	5	4	6

## Incomplete views

	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$
$s_1$	5	10	2
$s_2$	1	1	2
$s_4$	5	4	3

	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$
$s_1$	10	7	2
$s_3$	3	4	7

# Boost.SH





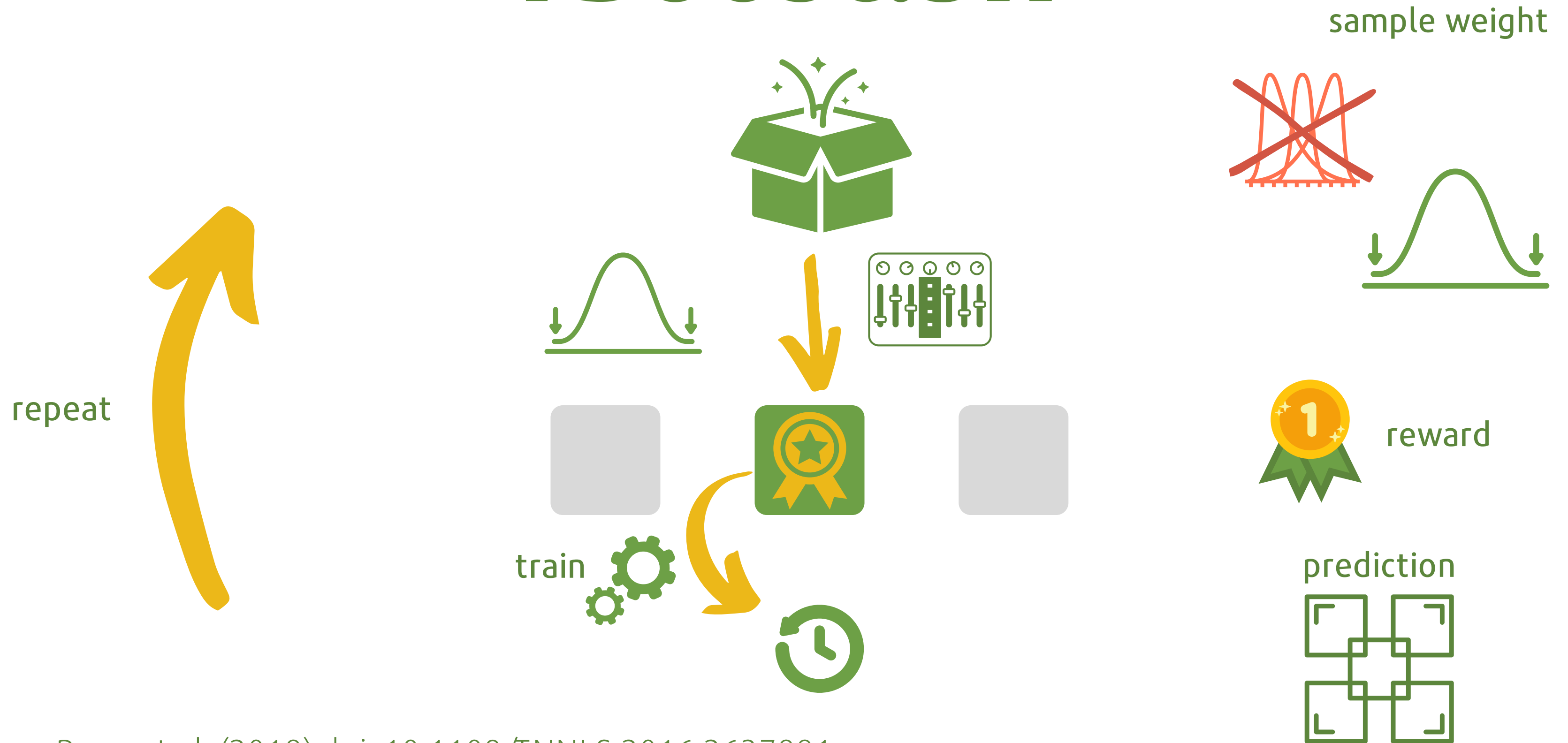


# Adversarial MAB

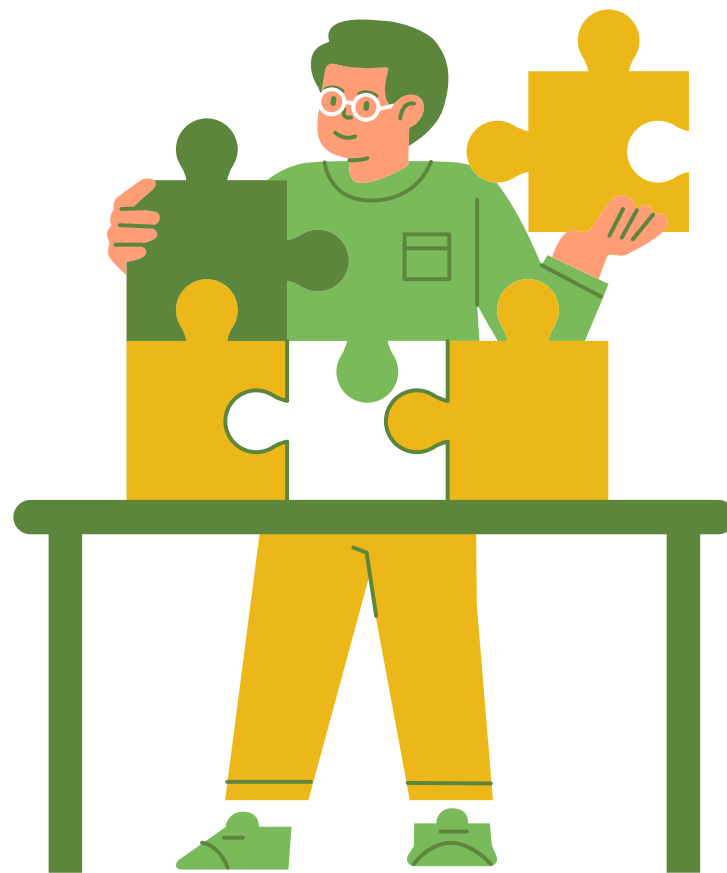
- Multiarmed bandits (MAB)
- Player vs  $k$  slot machines
- Maximize total reward
- Adversaries
- EXP3.P
- Very good in finding expected



# rBoost.SH



# ALGORITHM





# irBoost.SH

- Extension to rBoost.SH
- Incomplete views
- Multi-class setting
- Edge and prediction changes

## Complete views

	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$
$s_1$	5	10	2
$s_2$	1	1	2
$s_3$	7	4	1
$s_4$	5	4	3

	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$
$s_1$	10	7	2
$s_2$	1	2	2
$s_3$	3	4	7
$s_4$	5	4	6

## Incomplete views

	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$
$s_1$	5	10	2
$s_2$	1	1	2
$s_4$	5	4	3

	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$
$s_1$	10	7	2
$s_3$	3	4	7

# rBoost.SH

# irBoost.SH

$h_t(x_i)$  classifier's prediction

$y_i, h_t(x_i) \in \mathbb{N}$

$i$   
samples    time  $t$      $y_i, h_t(x_i) \in \{1, -1\}$

edge

edge

$$\theta_t = 2 \sum_{i \in \mathcal{N}} w_{t,i} y_i h_t(x_i)$$

$$\theta_t = 2 \sum_{i \in \mathcal{N}} w_i \left( \mathbb{I}[y_i = h_t(x_i)] - \frac{1}{2} \right)$$

edge weight

$$\alpha_t = \frac{1}{2} \ln \frac{1 + \theta_{t,j^*}}{1 - \theta_{t,j^*}} \quad \text{chosen view } j^*$$

reward

$$r_t = 1 - \sqrt{1 - \theta_t^2}$$

# rBoost.SH

weight update

$$w_{t+1,i} = \frac{w_{t,i}}{Z_t} \times e^{-\alpha_t y_i h_t(x_i)}$$

$Z_t$  → normalization factor

prediction

$$H(\mathbf{x}) = \text{sign} \left( \sum_t \alpha_t h_{t,*}(x_{i,*}) \right)$$

# irBoost.SH

weight update

$$w_{t+1,i} = \frac{w_{t,i}}{Z_t} \times e^{-2\alpha_t (\mathbb{I}[y_i = h_t(x_i)] - \frac{1}{2})}$$

prediction

$$H(\mathbf{x}) = \arg \max_c \frac{1}{Z} \left( \sum_t \alpha_t \hat{\mathbf{h}}_t(x_{i,*}) \right)$$

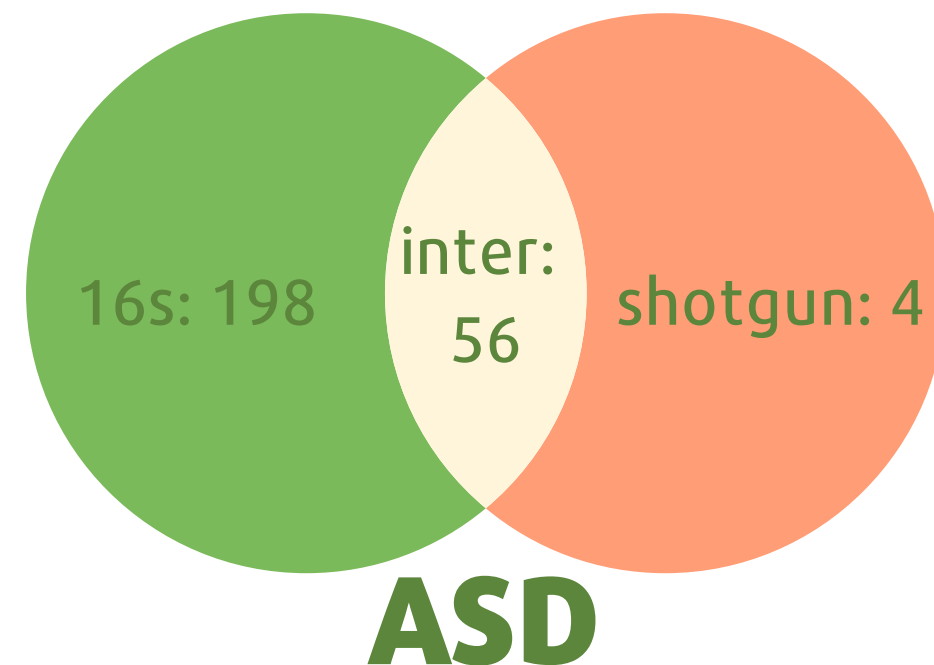
# RESULTS



# DATA

## ASD

- 2 incomplete views
- amplicon and shotgun data
- 2 classes (ASD vs noASD)

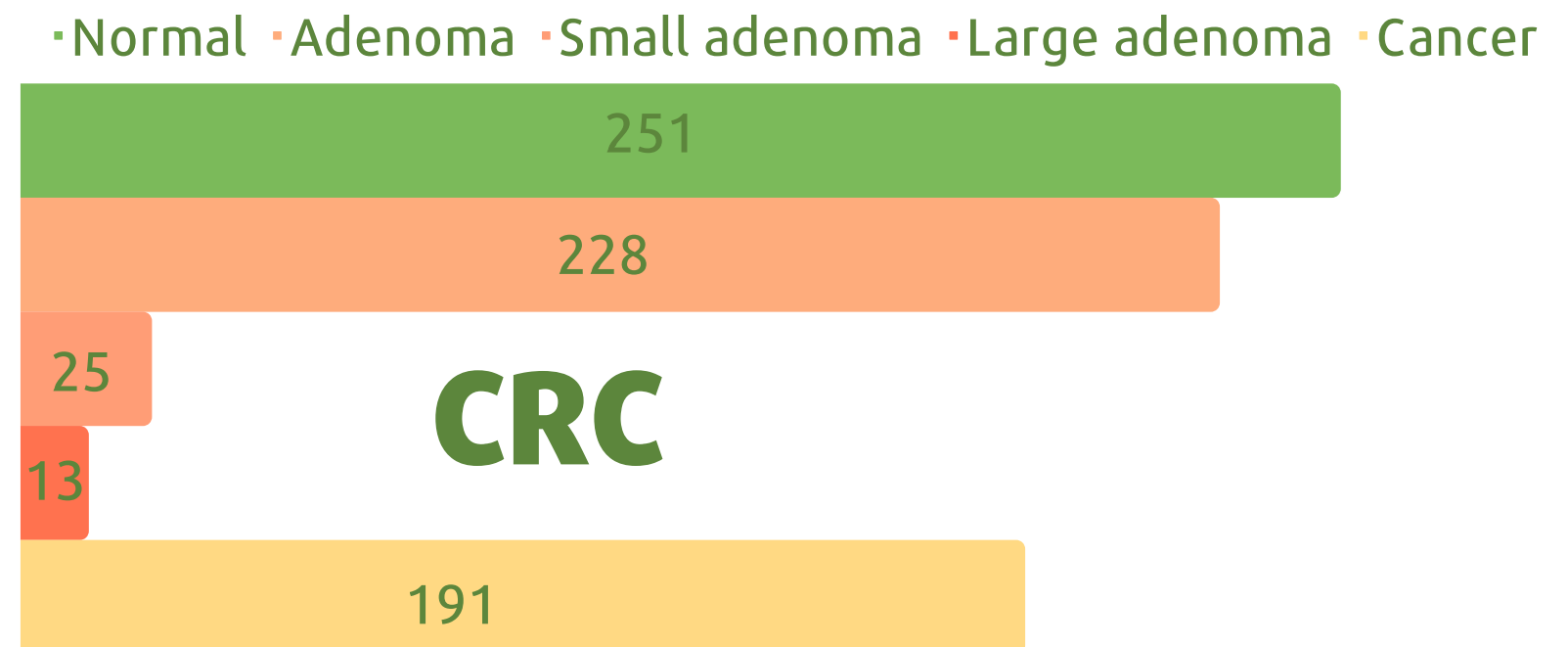


## ASD-16S

- 40 views from 16s sequences
- varying clustering similarity and filtration method

## CRC

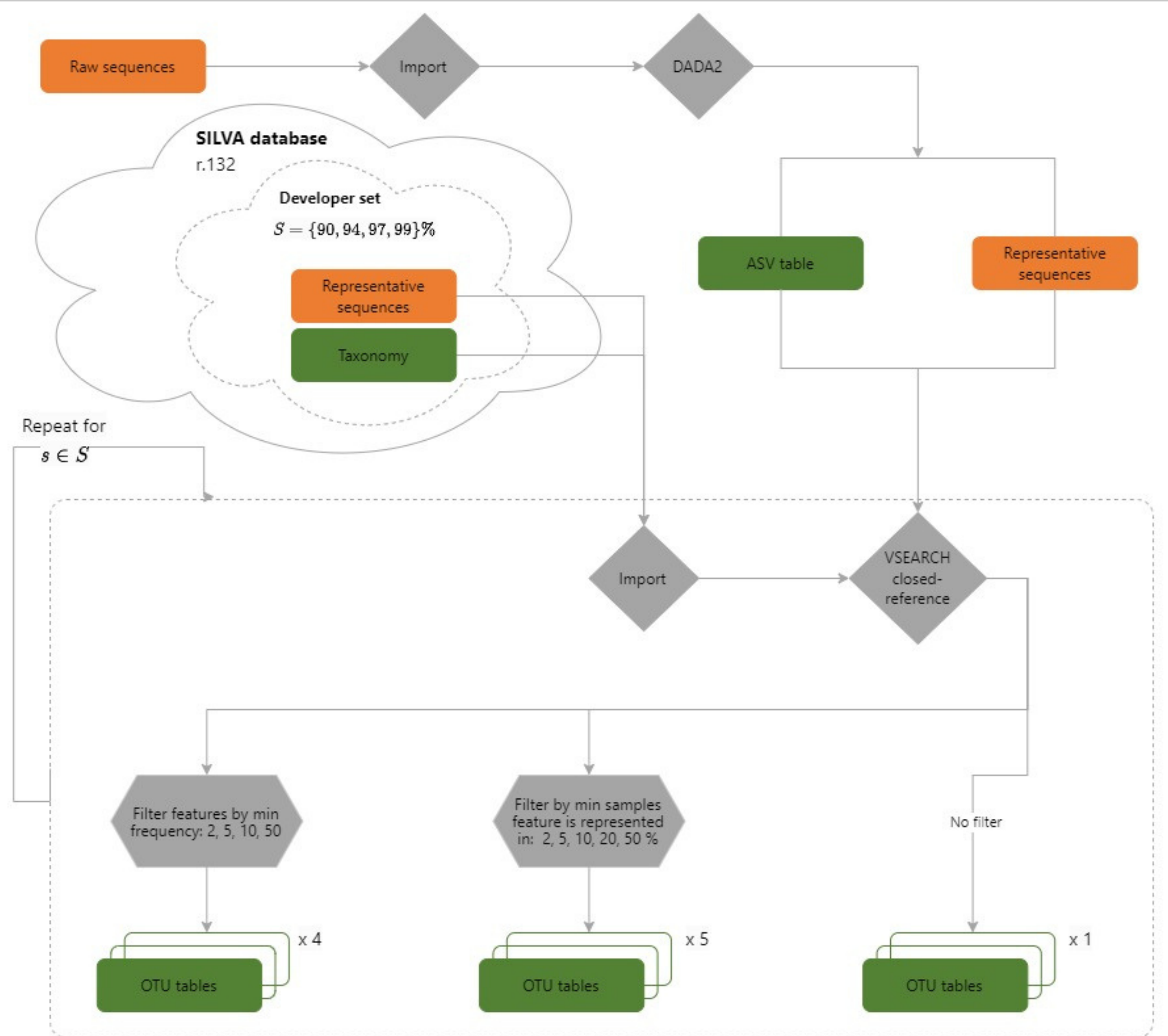
- 44 views from 16s sequences
- same pipeline like ASD-16s
- 3 classes (C, A, H)





# DATA PRE-PROCESSING

- Using QIIME 2
- SILVA taxonomy database
- Different similarity measures
- Filtering strategies:
  - filtering features by minimum frequency
  - filtering features by minimum samples in which feature must be represented
  - no filtering



# ASD EXPERIMENTS

## ASD-1

### ONLY 16S SAMPLES

- Single view
- Baseline multi-view
- irBoostSH

## ASD-2

### ONLY SHOTGUN SAMPLES

- Single view
- Baseline multi-view
- irBoostSH

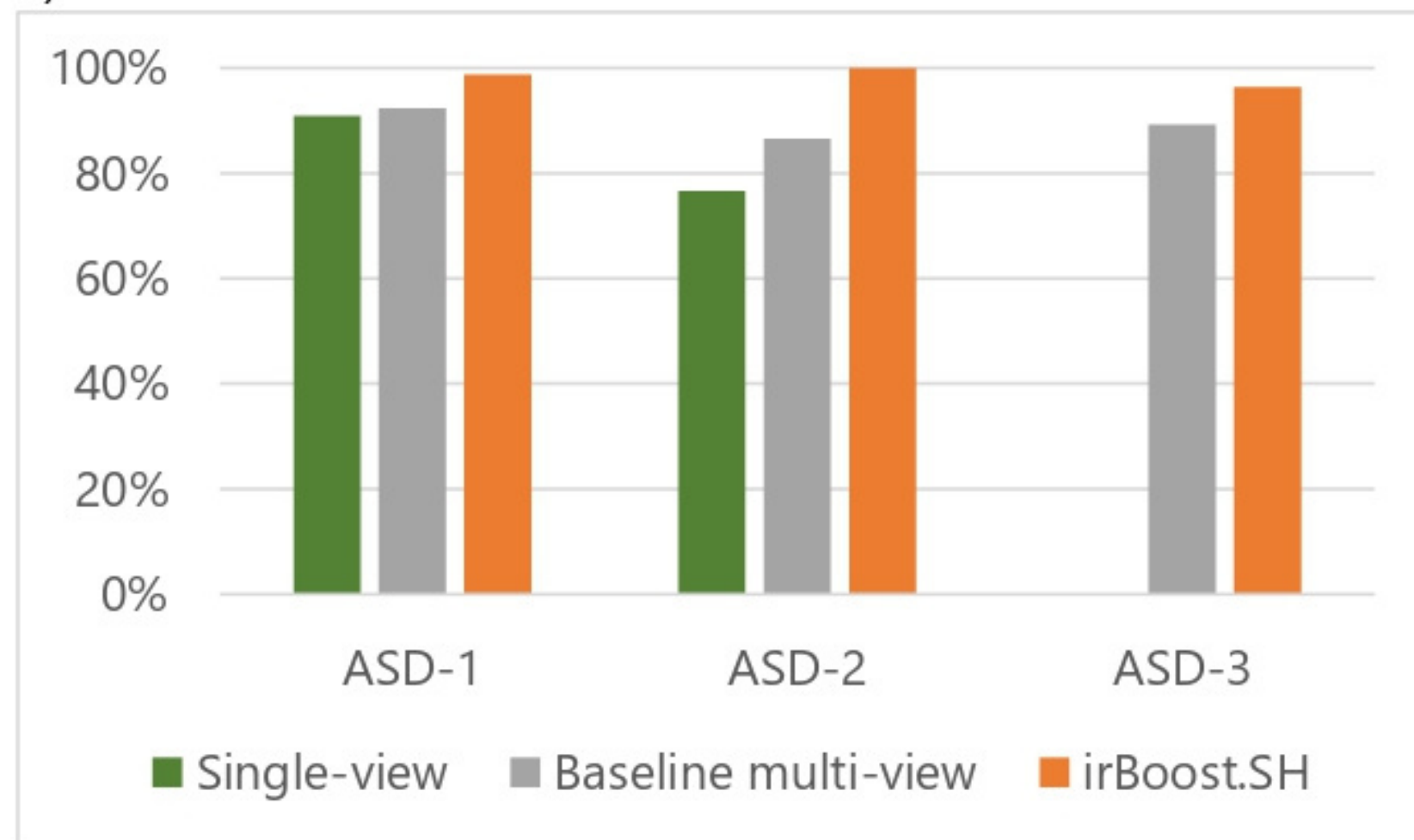
## ASD-3

### INTERSECTION SAMPLES

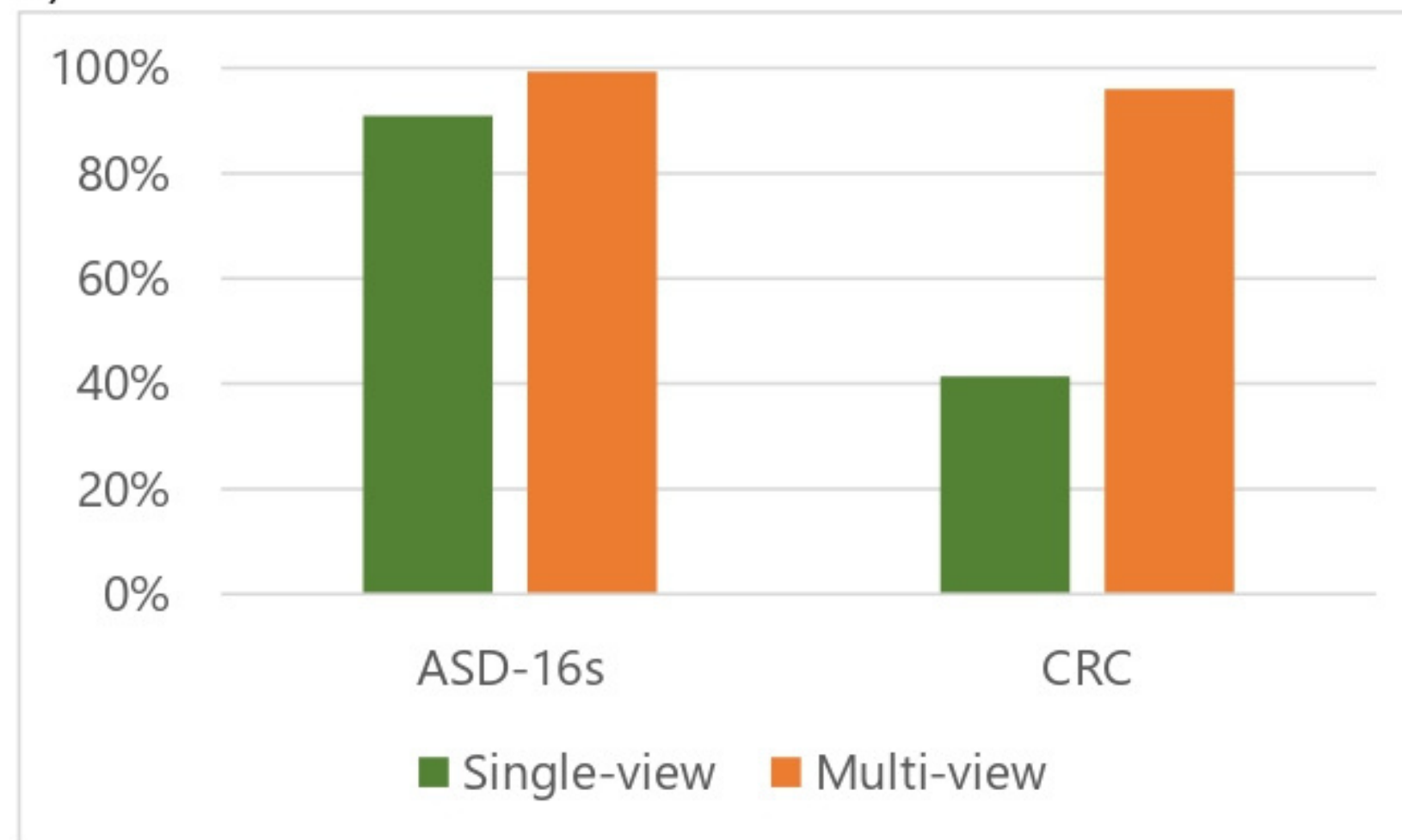
- Baseline multi-view
- irBoostSH

# RESULTS

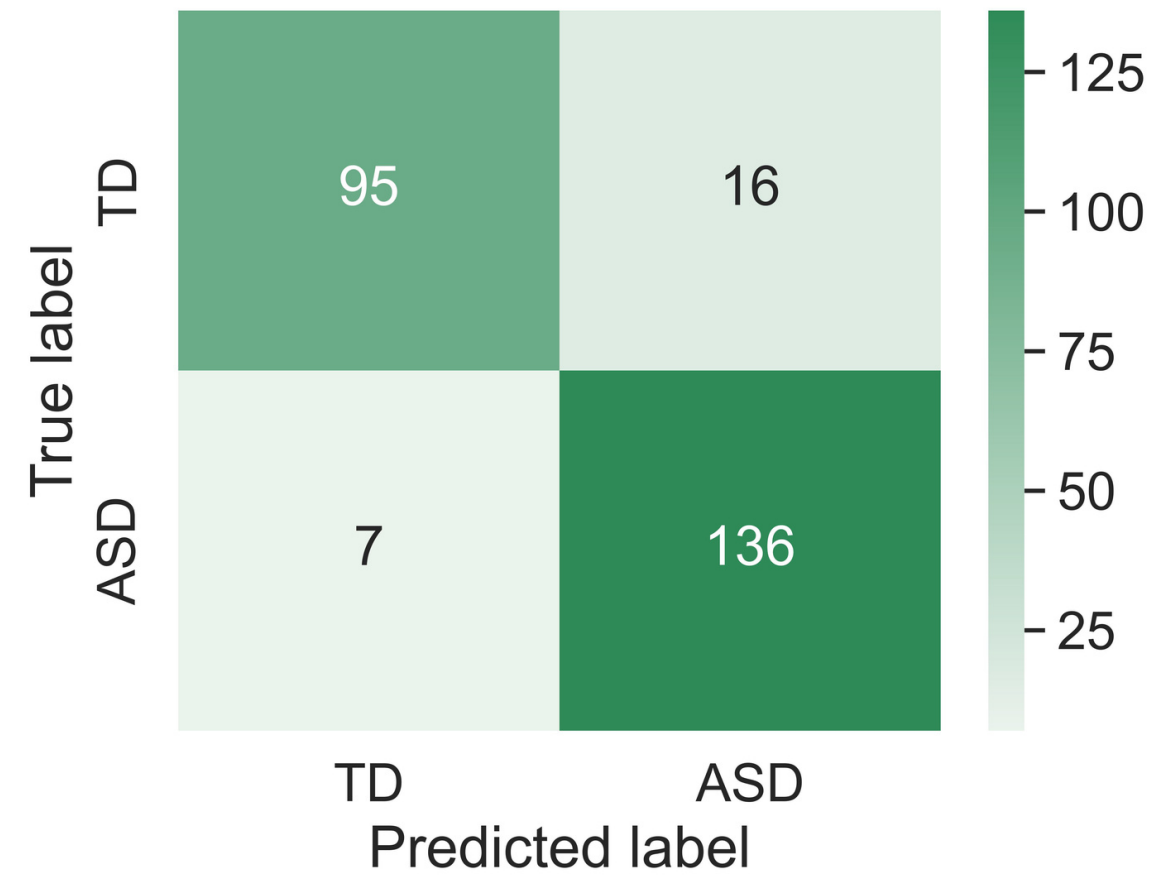
a)



b)

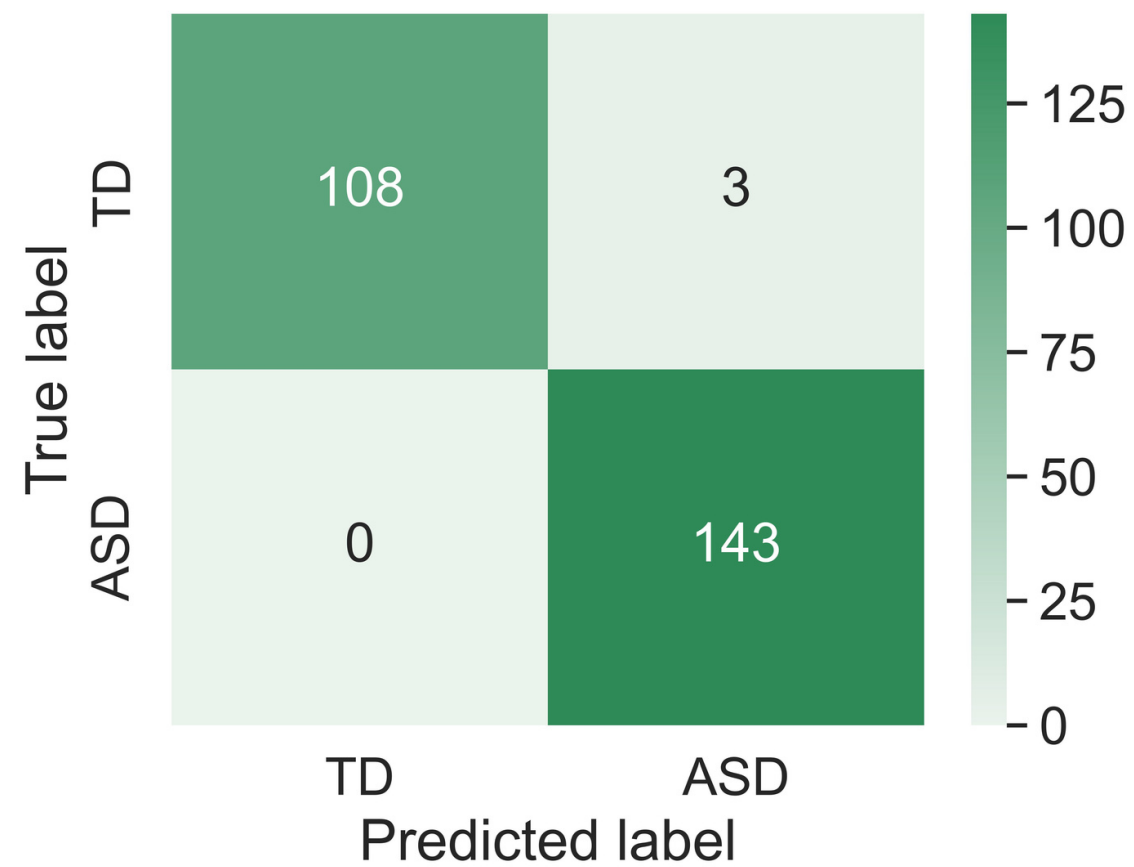


# Single-view

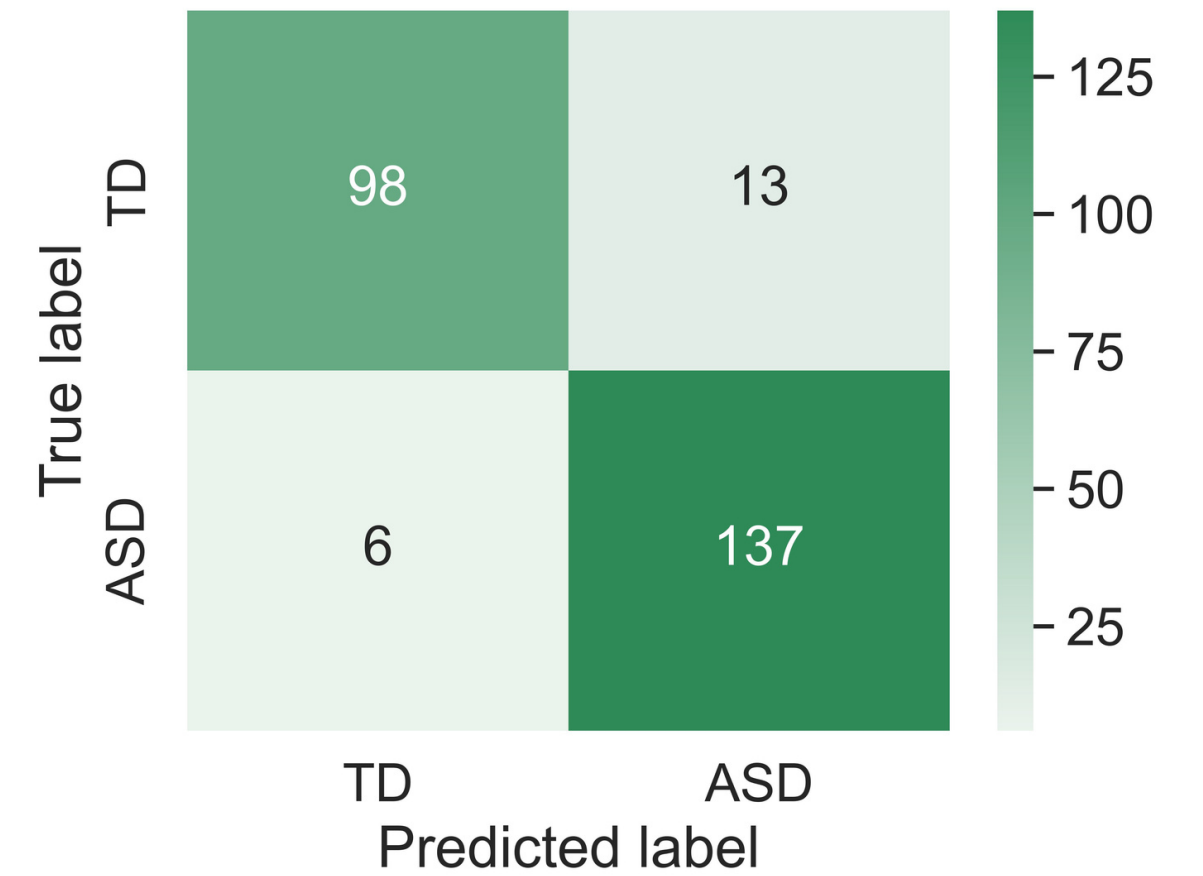


# ASD-1

## irBoost.SH

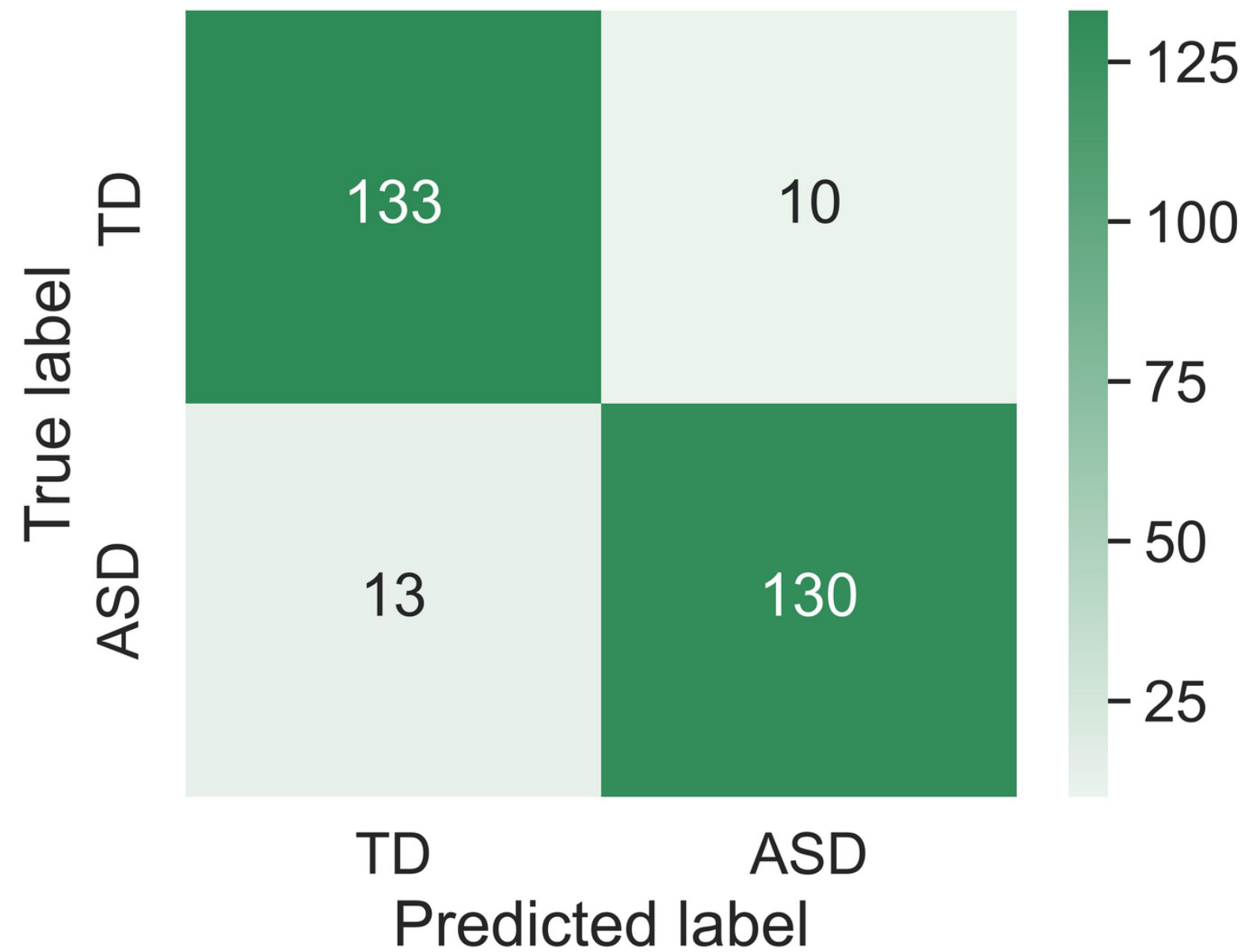


# Baseline

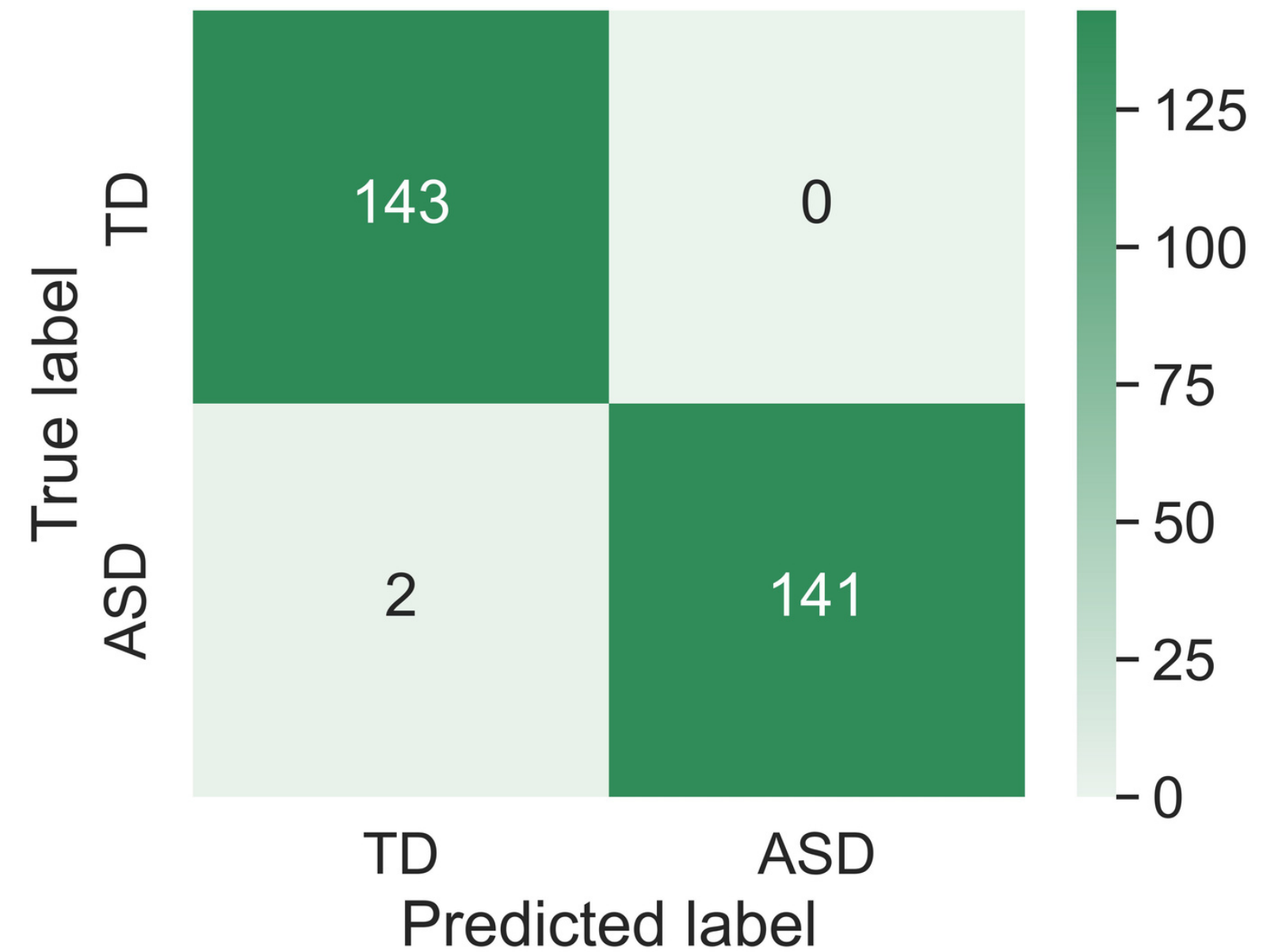


# ASD-16S

## Single-view



## irBoost.SH

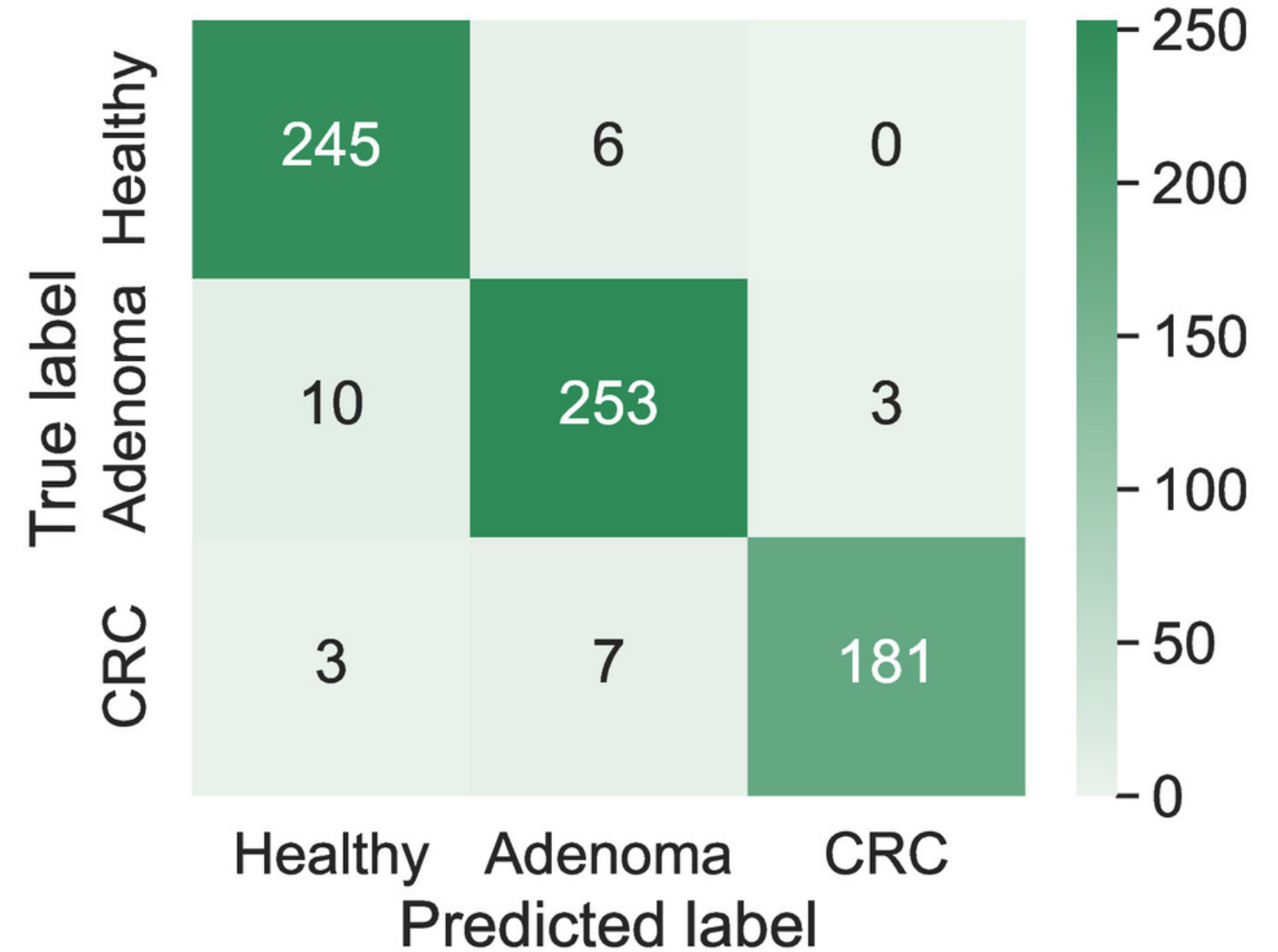


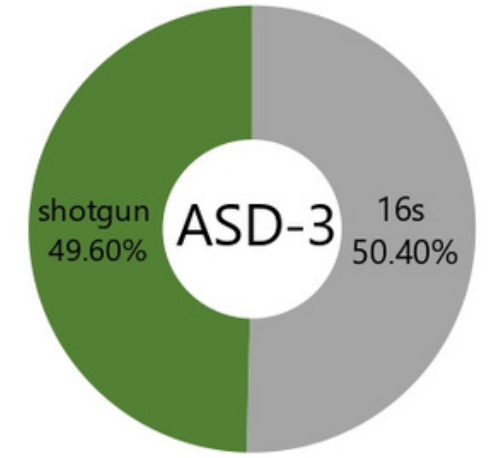
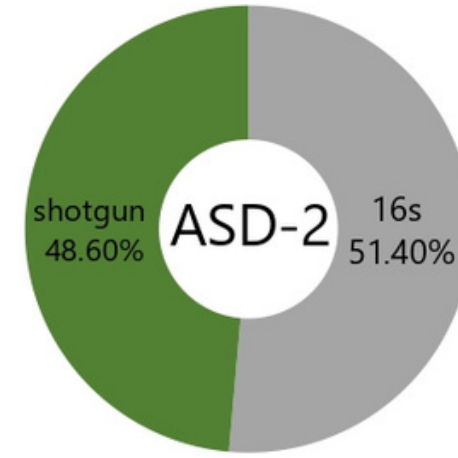
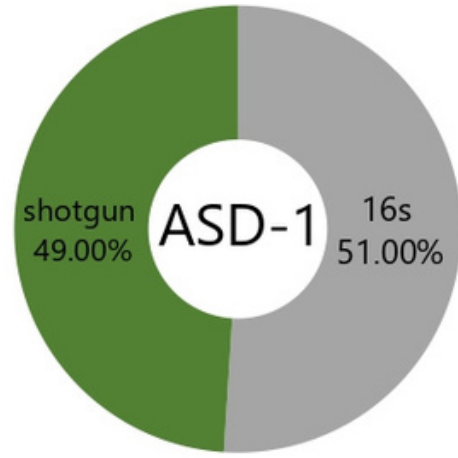
# CRC

## Single-view

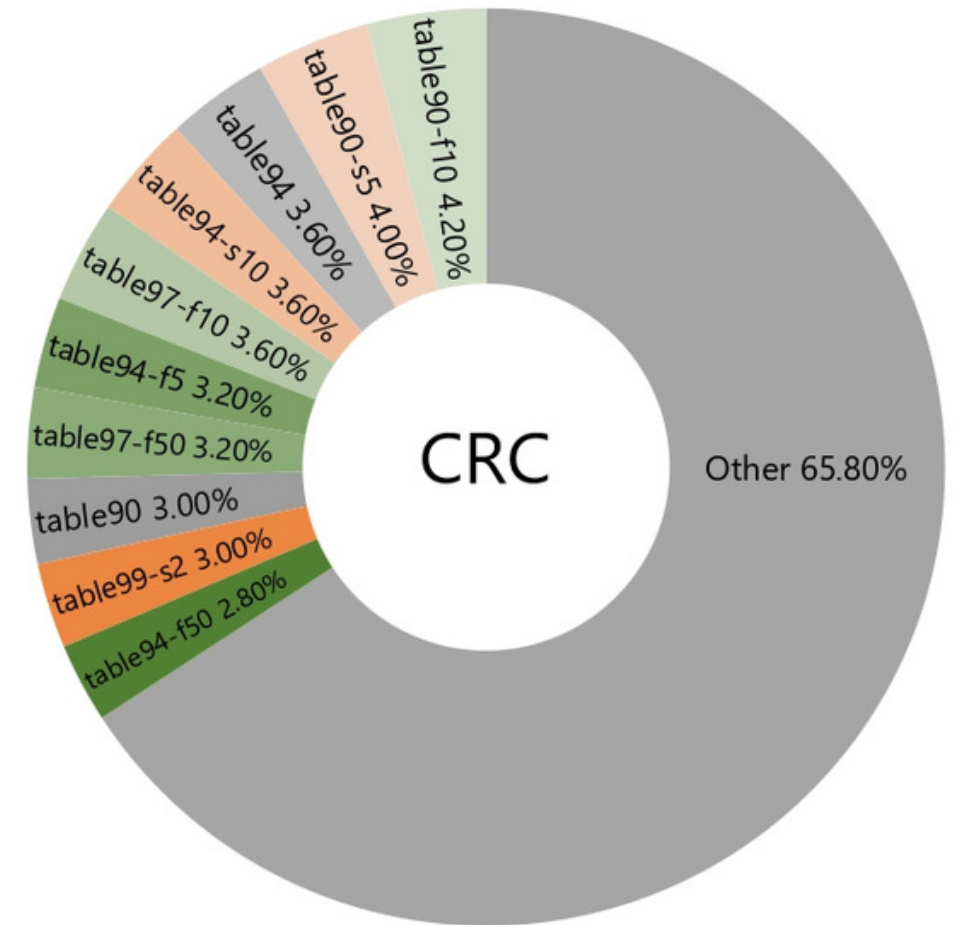
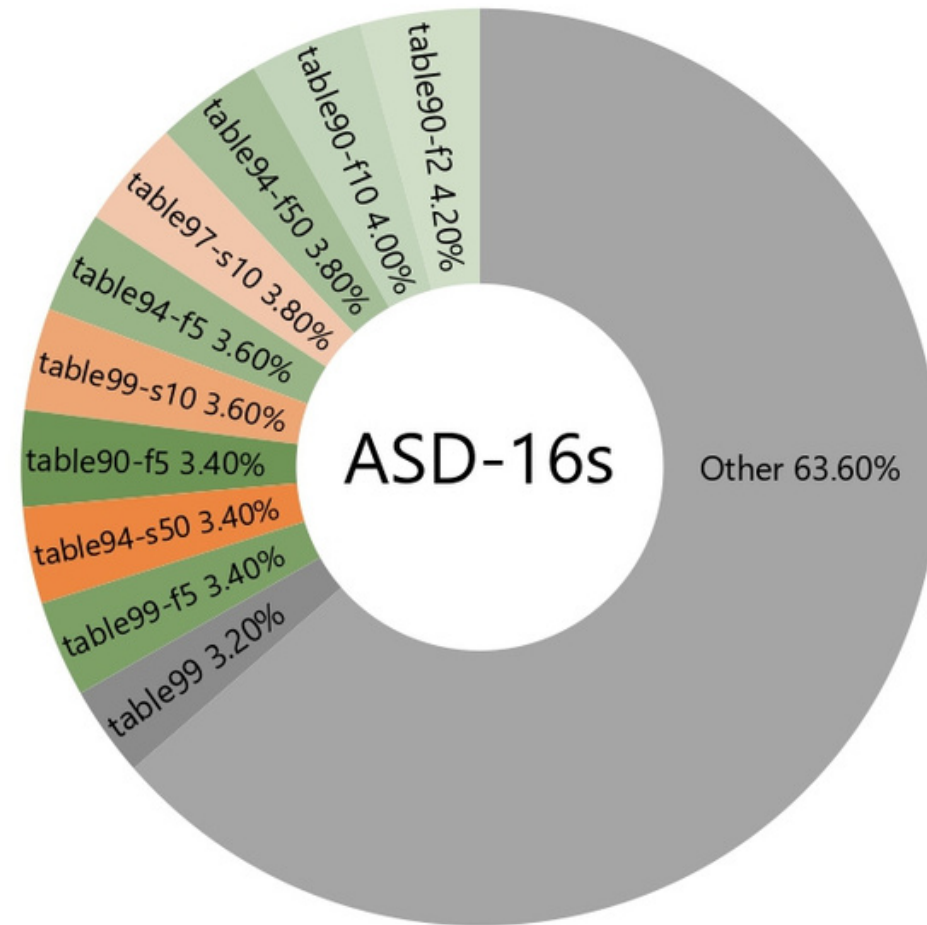


## irBoost.SH





# VIEW STATISTICS



# CONCLUSION

- Efficient multi-view framework
- Diversity necessary
- Extended:
  - multi-class case
  - view-based missing samples
- Learning rate?
- How much heterogeneous?
- Regression?
- Weight initialization?
- Semi-supervised?



# THANK YOU FOR LISTENING!

Q & A

