

Competition: Can I make a wish? Predicting the presence of meteors in images

Dr. Renato M. Silva, Prof. Dr. Tiago A. Almeida
Johannes V. Lochter

Intelligent and Distributed Systems Laboratory (LaSID)
Department of Computer Science (DComp)
Federal University of Sao Carlos (UFSCar)
Sorocaba – SP – Brasil

The logo for LaSID, featuring the text "LaSID" in a stylized font where the "i" has a small circle above it, all contained within a black rectangular box.

`renatoms@dt.fee.unicamp.br`, `talmeida@ufscar.br`,
`jlochter@acm.org`



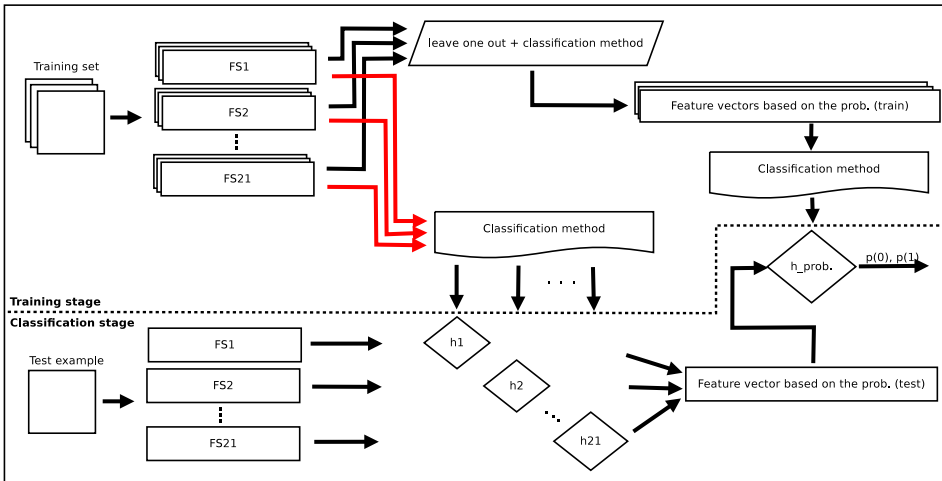
Introduction

- # training examples:
 - ▶ class 0 (non-meteor): 54
 - ▶ class 1 (meteor): 26
- Number of features: 3,451

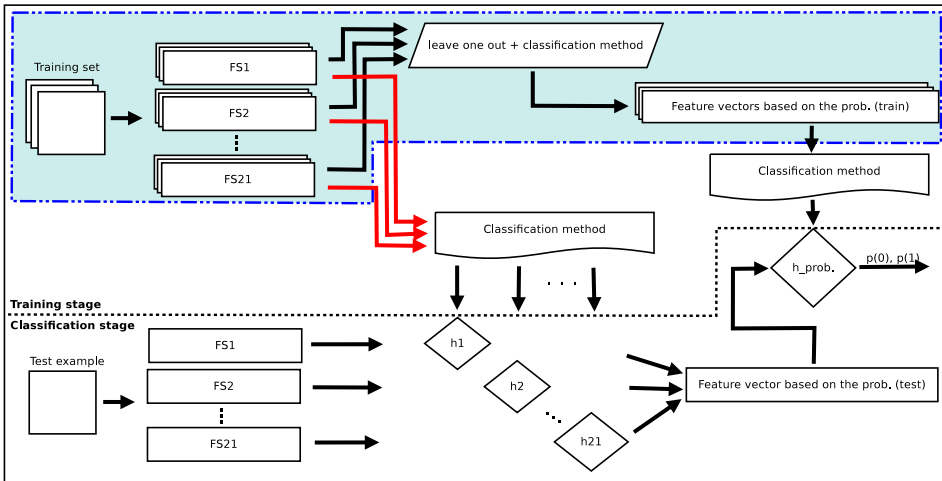
| Id | Feature set | # features |
|------|----------------------------|------------|
| FS1 | Auto Color Correlogram | 768 |
| FS2 | CEDD | 144 |
| FS3 | Color Histogram | 64 |
| FS4 | FCTH | 192 |
| FS5 | Fuzzy Histogram | 125 |
| FS6 | Fuzzy Opponent Histogram | 576 |
| FS7 | Gabor | 60 |
| FS8 | Haralick | 14 |
| FS9 | Histogram | 256 |
| FS10 | JCD | 168 |
| FS11 | Jpeg Coefficient Histogram | 192 |

| Id | Feature set | # features |
|------|--------------------------------------|------------|
| FS12 | Luminance Layout | 64 |
| FS13 | MPEG7 Color Layout | 33 |
| FS14 | MPEG7 Edge Histogram | 80 |
| FS15 | Mean Intensity Local Binary Patterns | 256 |
| FS16 | Mean Patch Intensity Histogram | 256 |
| FS17 | Moments | 4 |
| FS18 | Opponent Histogram | 64 |
| FS19 | PHOG | 40 |
| FS20 | Reference Color Similarity | 77 |
| FS21 | Tamura | 18 |

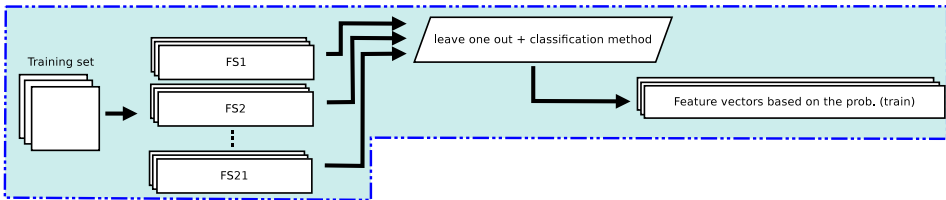
First selected submission



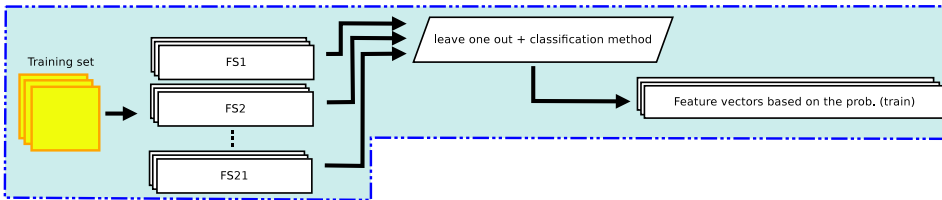
First selected submission



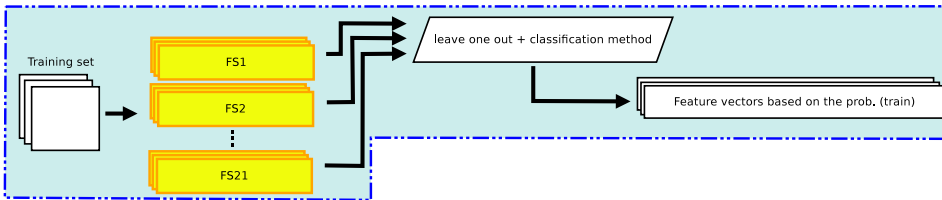
First selected submission



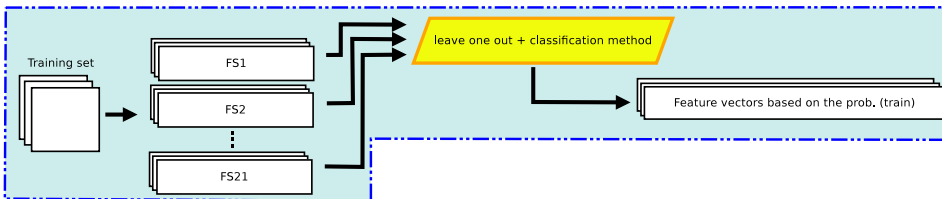
First selected submission



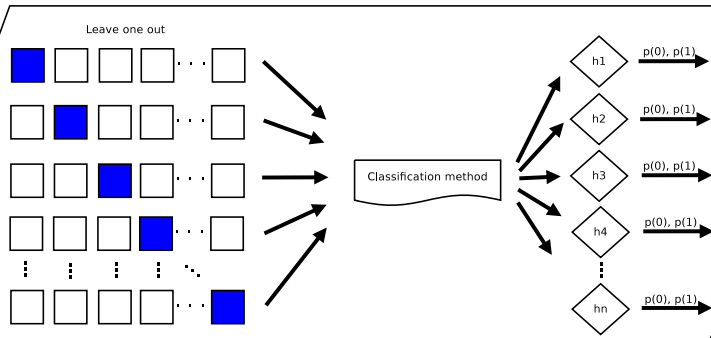
First selected submission



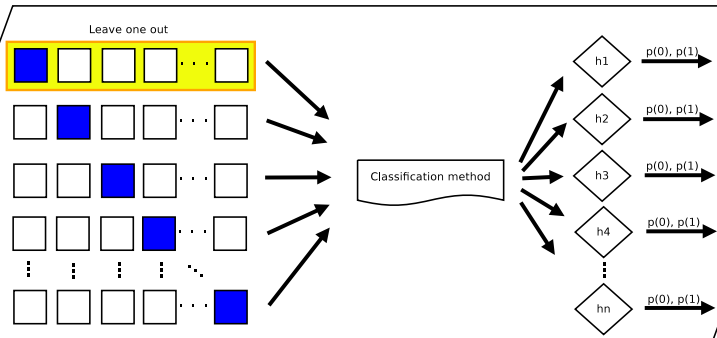
First selected submission



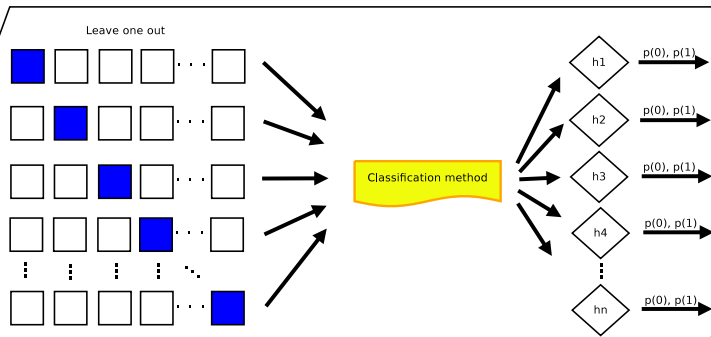
First selected submission



First selected submission

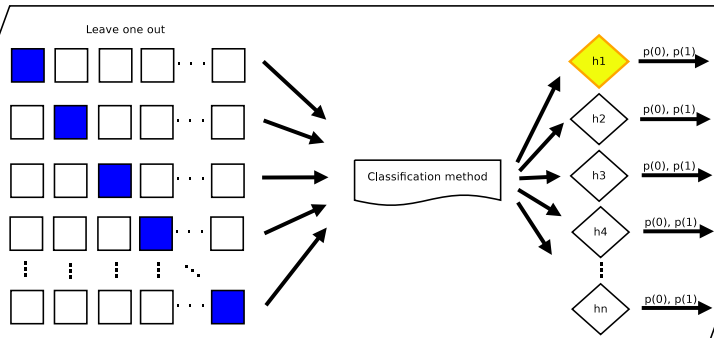


First selected submission

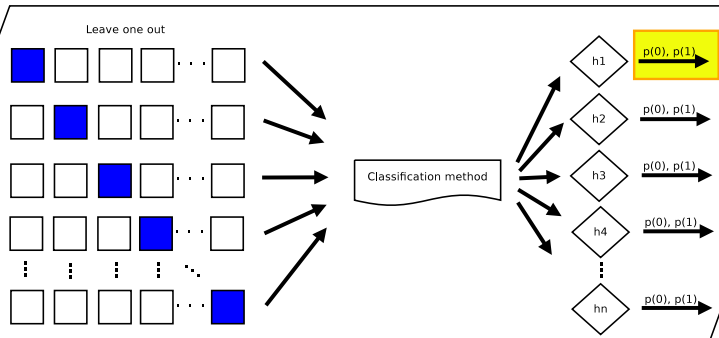


- classification method: Logistic Regression
- class balancing: SMOTE (Synthetic Minority Over-sampling Technique)

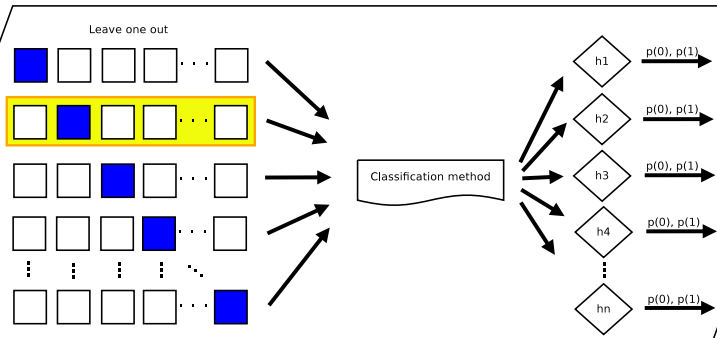
First selected submission



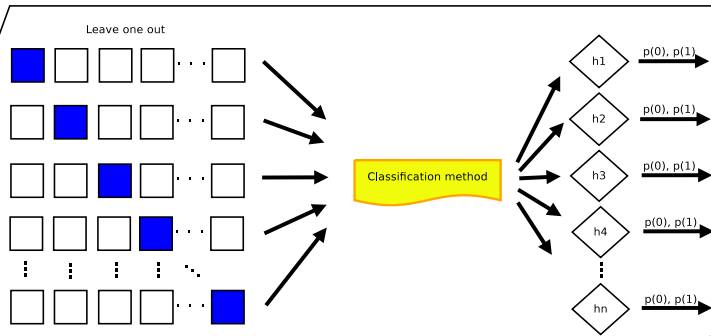
First selected submission



First selected submission

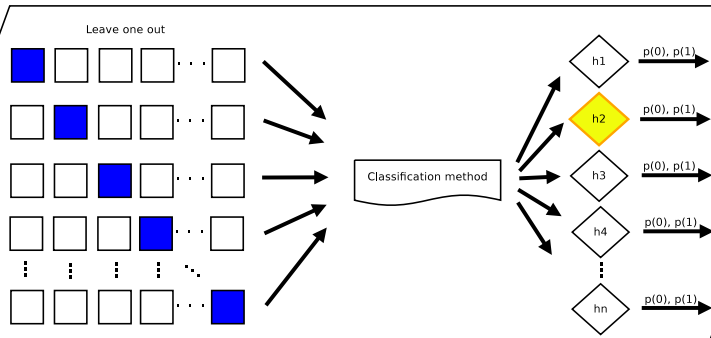


First selected submission

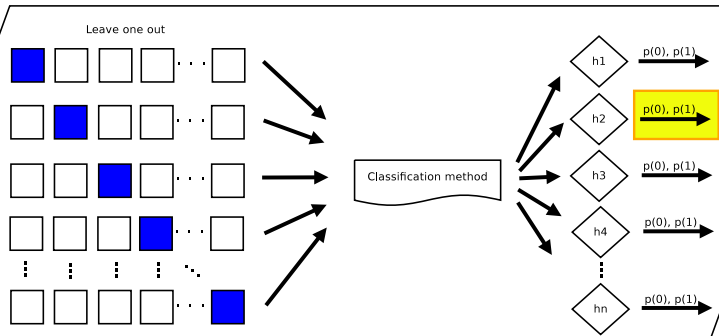


- classification method: Logistic Regression
- class balancing: SMOTE (Synthetic Minority Over-sampling Technique)

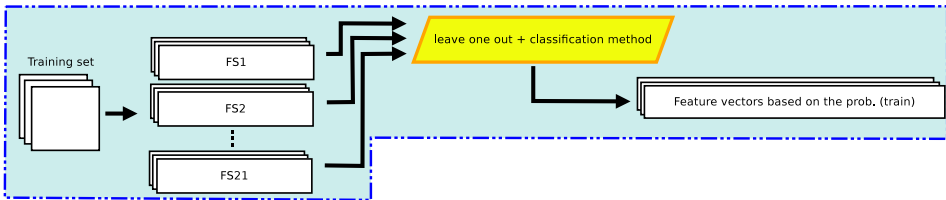
First selected submission



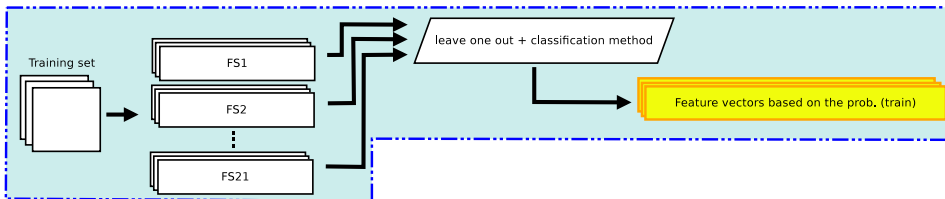
First selected submission



First selected submission

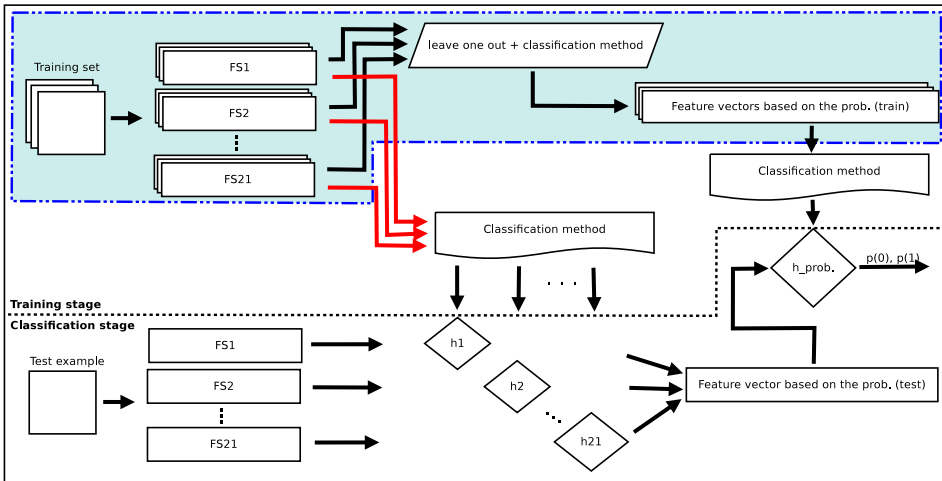


First selected submission

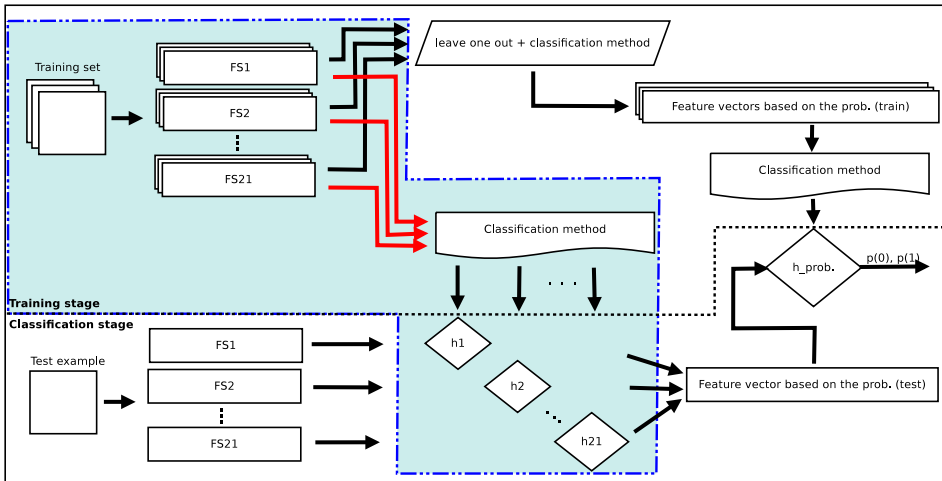


| Id | FS1 | FS2 | FS3 | ... | FS21 |
|-----|---------------------------------|---------------------------------|---------------------------------|-----|----------------------------------|
| 1 | $\text{Prob}_{\text{Id}=1,1}$ | $\text{Prob}_{\text{Id}=1,2}$ | $\text{Prob}_{\text{Id}=1,3}$ | ... | $\text{Prob}_{\text{Id}=1,21}$ |
| 2 | $\text{Prob}_{\text{Id}=2,1}$ | $\text{Prob}_{\text{Id}=2,2}$ | $\text{Prob}_{\text{Id}=2,3}$ | ... | $\text{Prob}_{\text{Id}=2,21}$ |
| 5 | $\text{Prob}_{\text{Id}=5,1}$ | $\text{Prob}_{\text{Id}=5,2}$ | $\text{Prob}_{\text{Id}=5,3}$ | ... | $\text{Prob}_{\text{Id}=5,21}$ |
| 8 | $\text{Prob}_{\text{Id}=8,1}$ | $\text{Prob}_{\text{Id}=8,2}$ | $\text{Prob}_{\text{Id}=8,3}$ | ... | $\text{Prob}_{\text{Id}=8,21}$ |
| 10 | $\text{Prob}_{\text{Id}=10,1}$ | $\text{Prob}_{\text{Id}=10,2}$ | $\text{Prob}_{\text{Id}=10,3}$ | ... | $\text{Prob}_{\text{Id}=10,21}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 122 | $\text{Prob}_{\text{Id}=122,1}$ | $\text{Prob}_{\text{Id}=122,2}$ | $\text{Prob}_{\text{Id}=122,3}$ | ... | $\text{Prob}_{\text{Id}=122,21}$ |

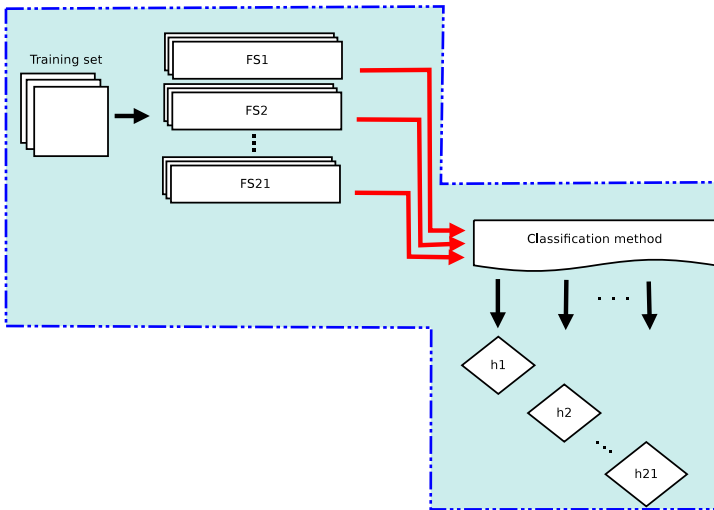
First selected submission



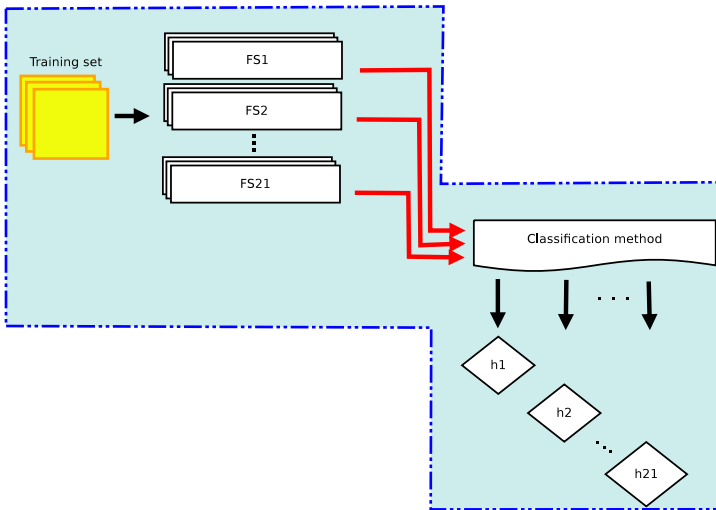
First selected submission



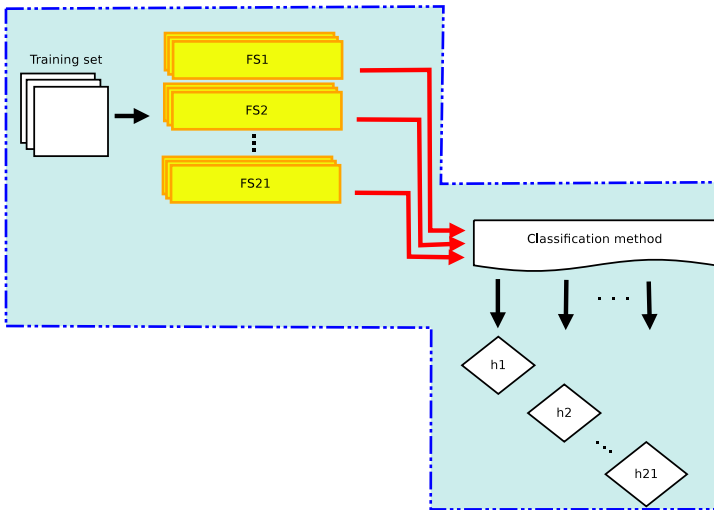
First selected submission



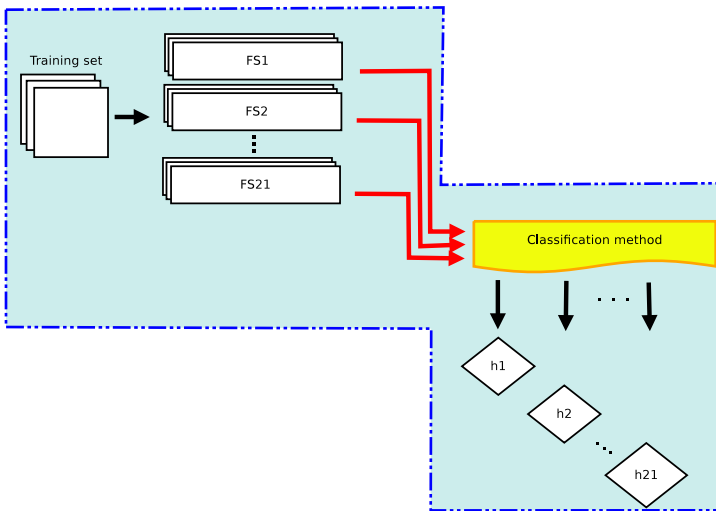
First selected submission



First selected submission

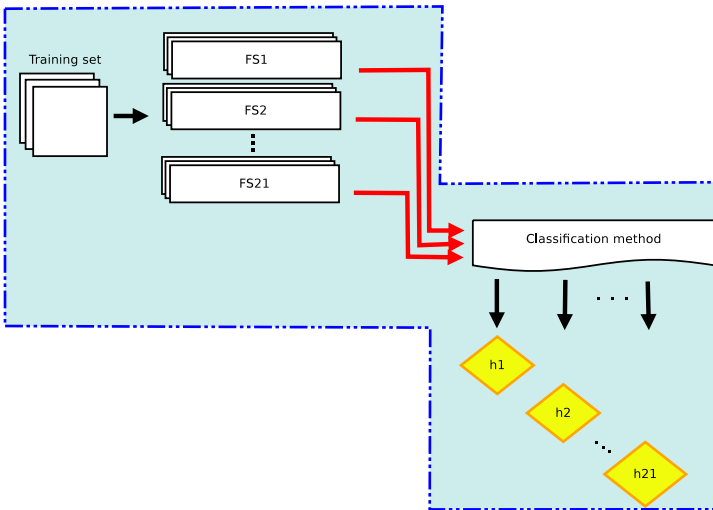


First selected submission

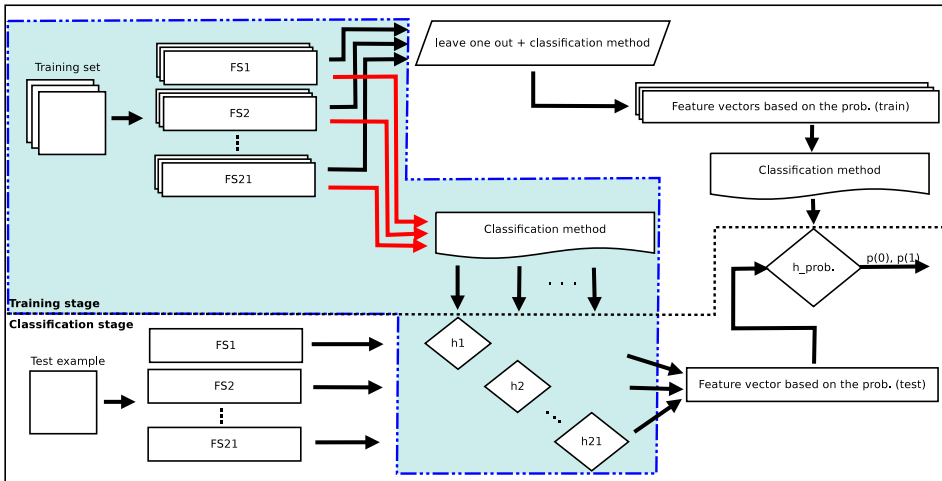


- classification method: Logistic Regression
- class balancing: SMOTE (Synthetic Minority Over-sampling Technique)

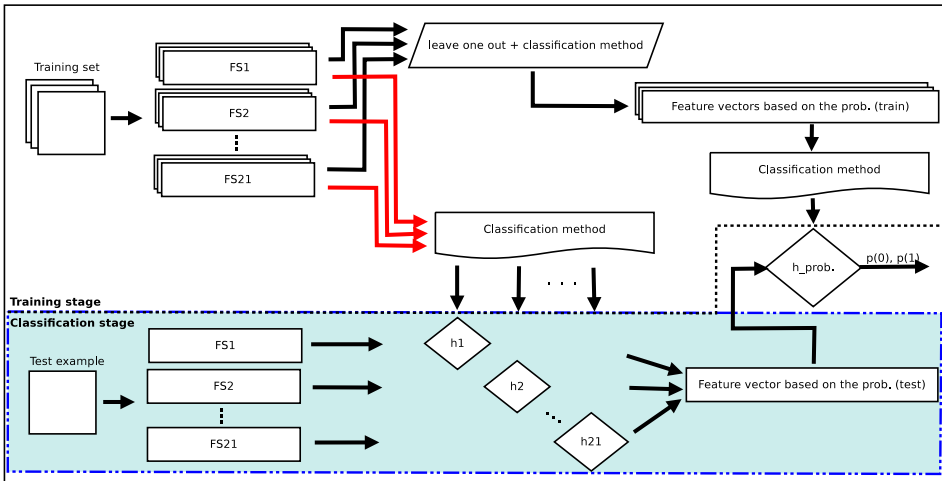
First selected submission



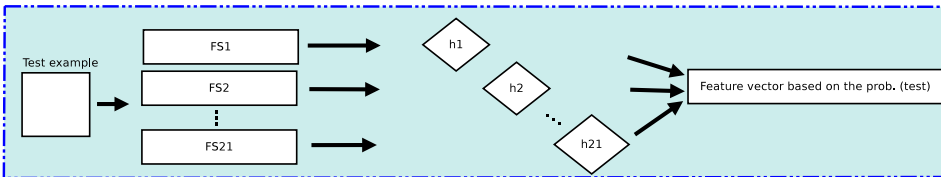
First selected submission



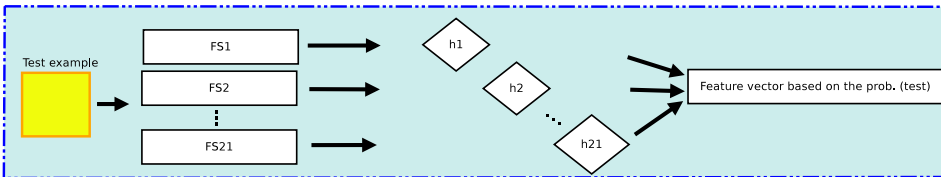
First selected submission



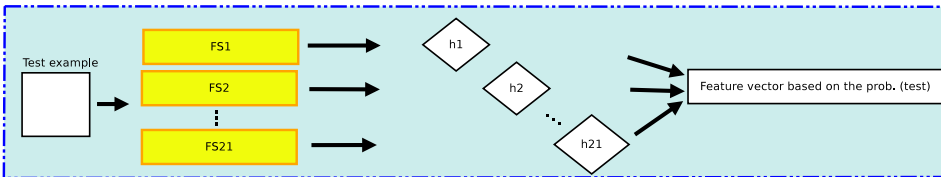
First selected submission



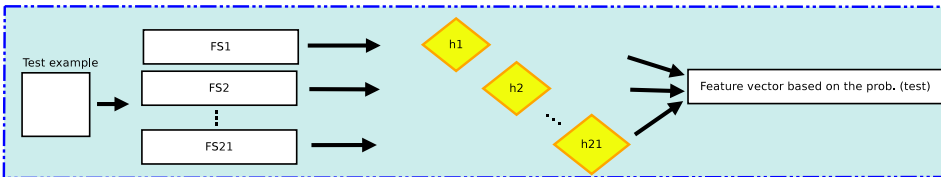
First selected submission



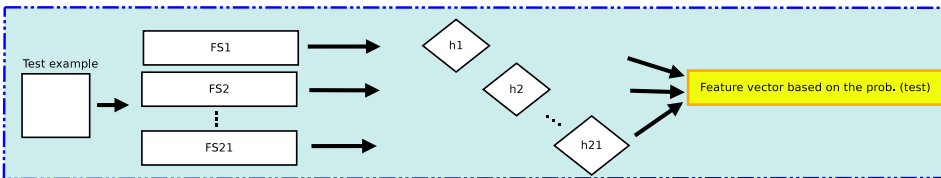
First selected submission



First selected submission

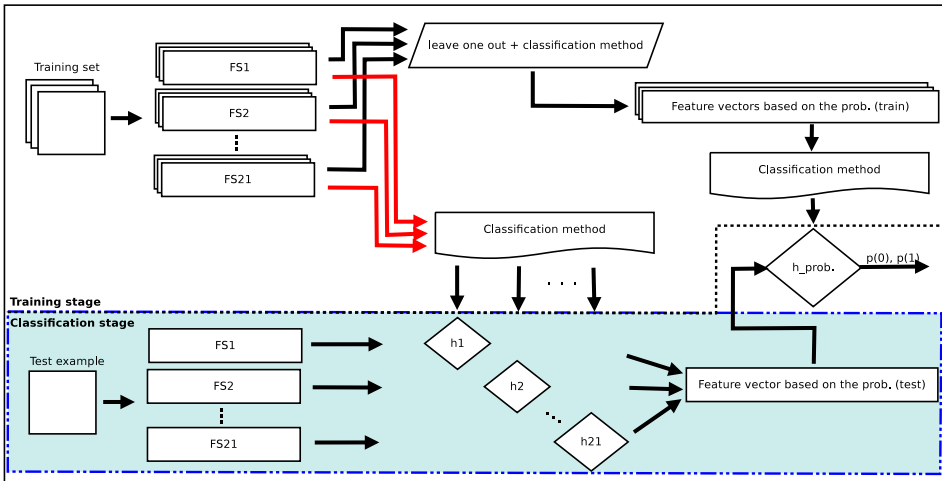


First selected submission

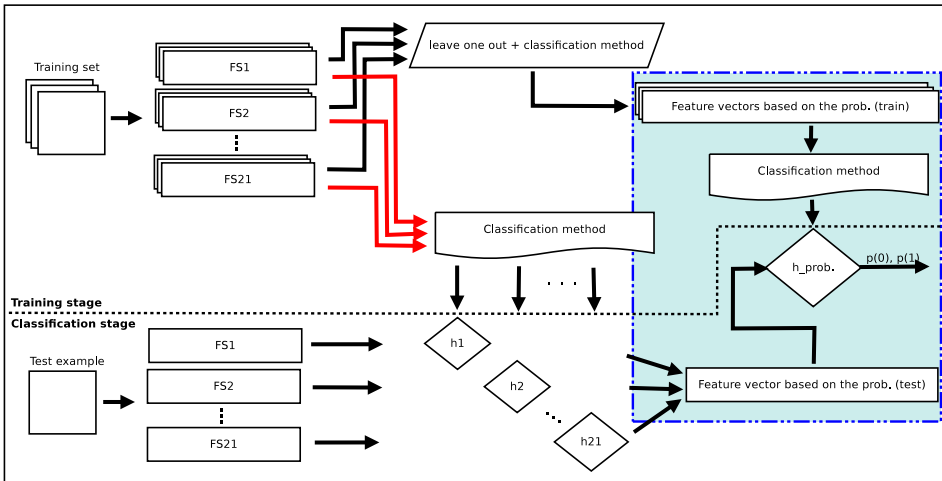


| Id | FS1 | FS2 | FS3 | ... | FS21 |
|-----|-------------------------------|-------------------------------|-------------------------------|-----|--------------------------------|
| i | $\text{Prob}_{\text{Id}=i,1}$ | $\text{Prob}_{\text{Id}=i,2}$ | $\text{Prob}_{\text{Id}=i,3}$ | ... | $\text{Prob}_{\text{Id}=i,21}$ |

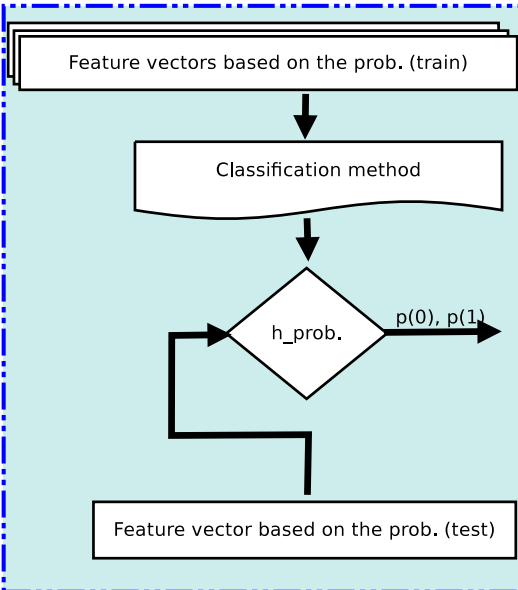
First selected submission



First selected submission



First selected submission



First selected submission

Feature vectors based on the prob. (train)

Classification method

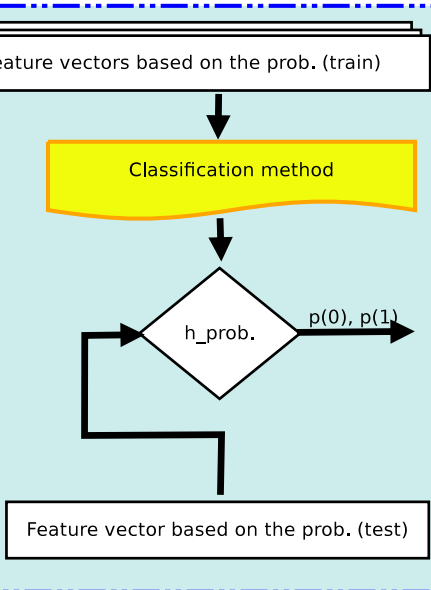
$h_{\text{prob.}}$

$p(0), p(1)$

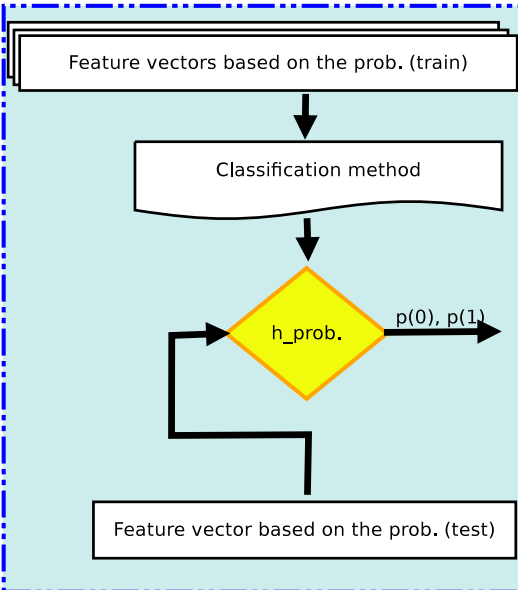
Feature vector based on the prob. (test)

First selected submission

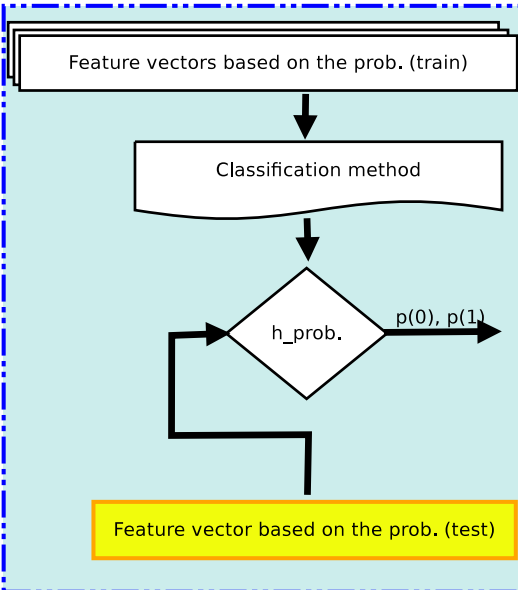
- classification method: Logistic Regression
- class balancing: SMOTE (Synthetic Minority Over-sampling Technique)



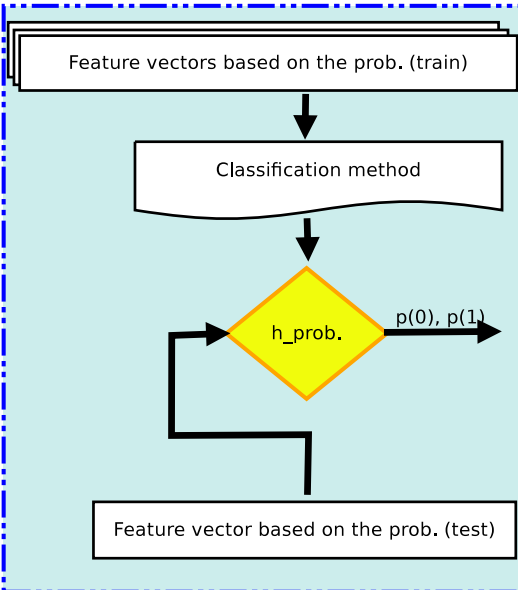
First selected submission



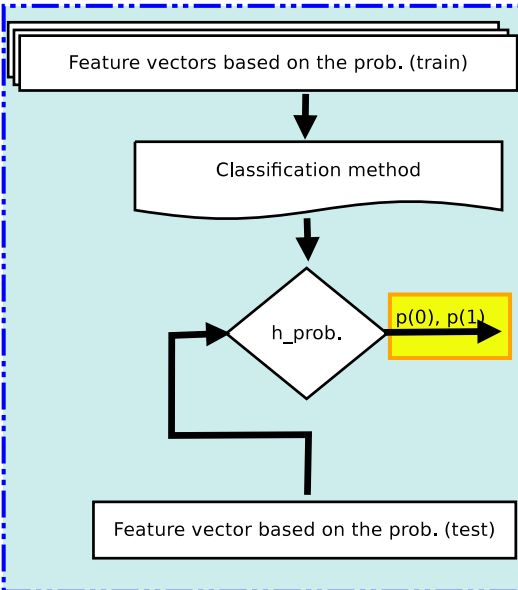
First selected submission



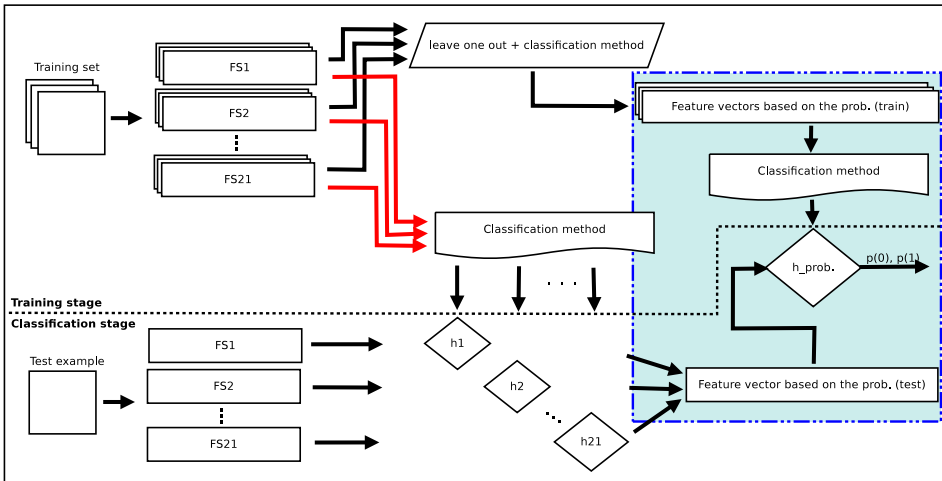
First selected submission



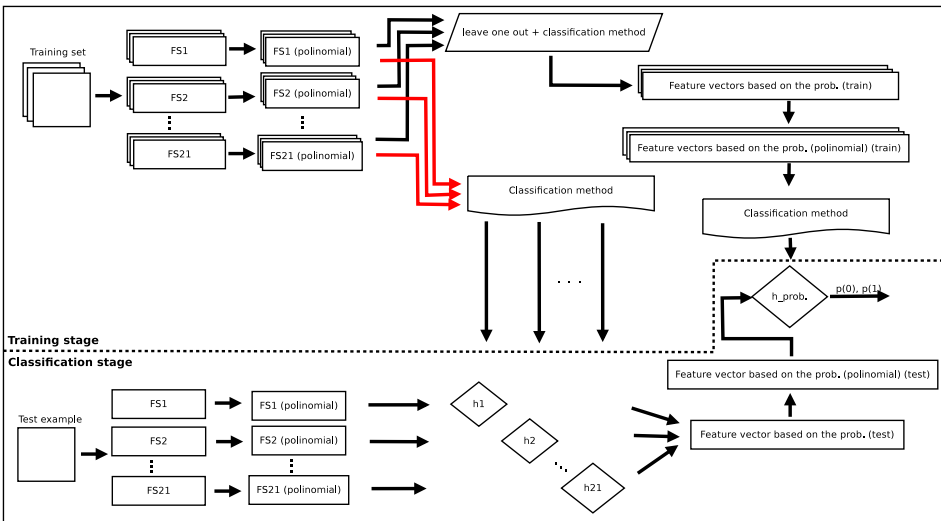
First selected submission



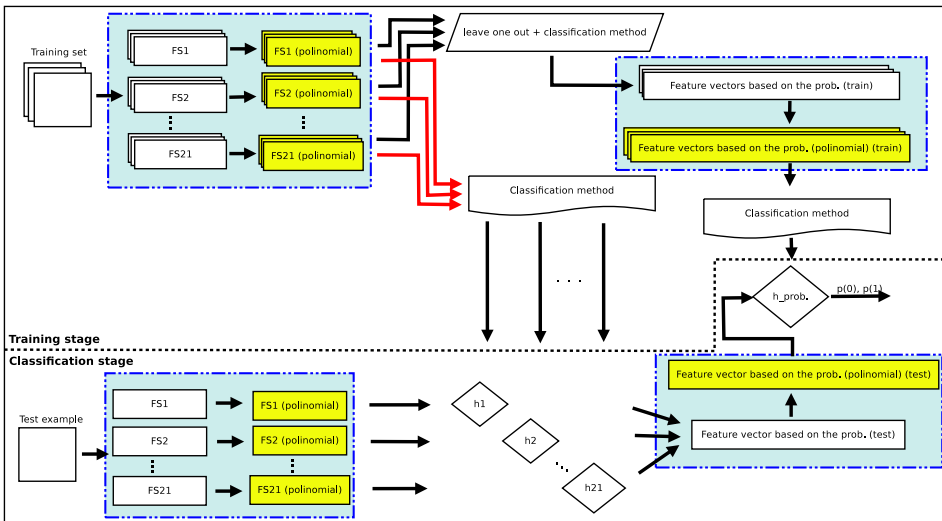
First selected submission



Second selected submission



Second selected submission



- ❑ Python 3.5
- ❑ Logistic Regression: `scikit-learn` library
- ❑ Polynomial features: `scikit-learn` library
- ❑ SMOTE: `imbalanced-learn` library

□ URL: <https://github.com/renatoms88/KDDBR>

renatoms88 / KDDBR

Unwatch 1 Star 0 Fork 0

<> Code Issues 0 Pull requests 0 Projects 0 Wiki Settings Insights

No description, website, or topics provided. Edit

Add topics

3 commits

1 branch

0 releases

1 contributor

Branch: master New pull request

Create new file Upload files Find file Clone or download

| | | |
|---------------------------|------------------------------------|----------------|
| renatoms88 update | Latest commit be3aa5 6 minutes ago | |
| dataset/DatasetPerFeature | update | 21 minutes ago |
| README.md | update | 6 minutes ago |
| main.py | update | 21 minutes ago |

README.md

Competition: Can I make a wish? Predicting the presence of meteors in images

Here I present the implementation code of my team's approach for solving the 1st KDD-BR (Brazilian Knowledge Discovery in Databases) competition hosted on Kaggle.

Competition: Can I make a wish? Predicting the presence of meteors in images

Dr. Renato M. Silva, Prof. Dr. Tiago A. Almeida
Johannes V. Lochter

Intelligent and Distributed Systems Laboratory (LaSID)
Department of Computer Science (DComp)
Federal University of Sao Carlos (UFSCar)
Sorocaba – SP – Brasil

The logo for LaSID, featuring the text "LaSID" in a stylized font where the "i" has a dot that looks like a small circle.

`renatoms@dt.fee.unicamp.br`, `talmeida@ufscar.br`,
`jlochter@acm.org`

