How to learn to automatically analyze MRIs in less than 6 months

Pablo Pérez Sánchez

PyData

DATALAB
MEDIALAB USAL

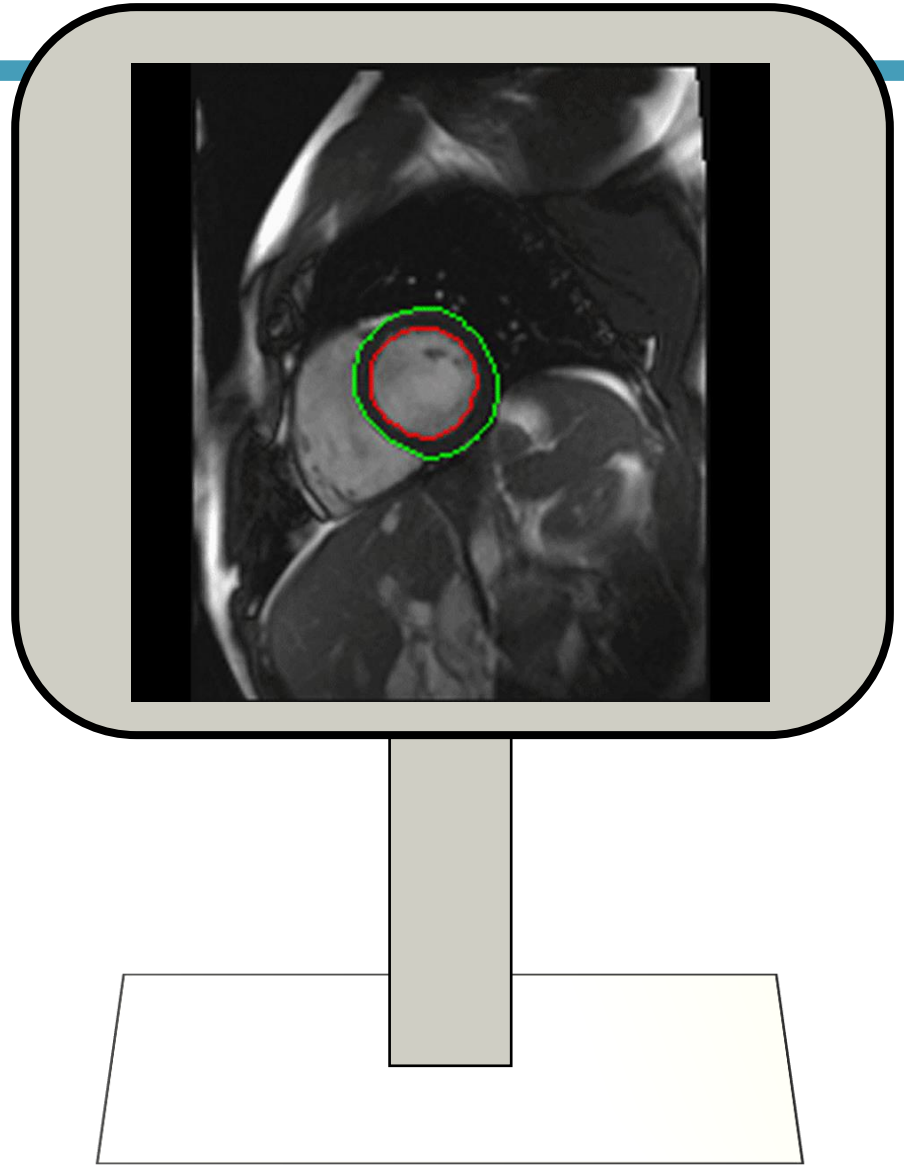
INTRODUCTION AND MOTIVATION

30 %

Deaths by heart diseases

Nowadays, doctors use Magnetic Resonance Imaging in order to measure the volume of heart chambers

Here we can observe how a MRI can be seen on a screen



PyData

DATALAB
MEDIALAB USAL

In 2014

8 – 10

MRIs per month*

Nowadays

210 – 240

MRIs per month*

* Data from Salamanca

PyData

DATALAB
MEDIALAB USAL

The MRIs provide valuable information, but its interpretation isn't automated and it takes some time. Because of this some problems arise.

PROBLEMS OF THE TRADITIONAL METHOD

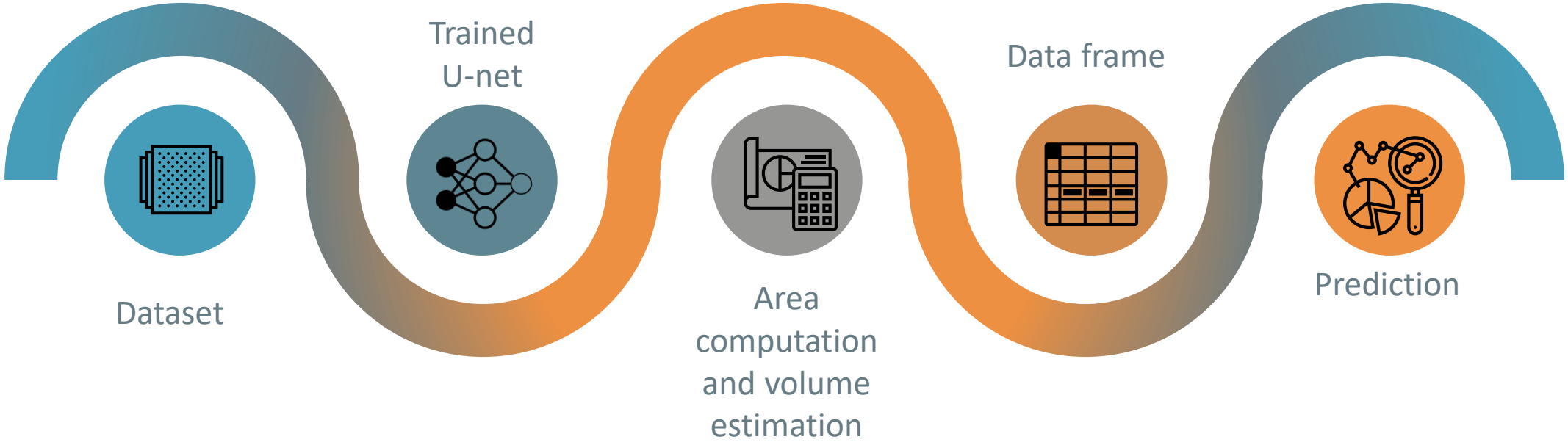
SEGMENTATION NOT VERY ACCURATE ON RIGHT VENTRICLE

DIRECT INTERVENTION OF A PHYSICIAN IN THE SEGMENTATION PROCESS

LESS TIME IS SPENT ON EACH PATIENT (~ 5 MIN.)

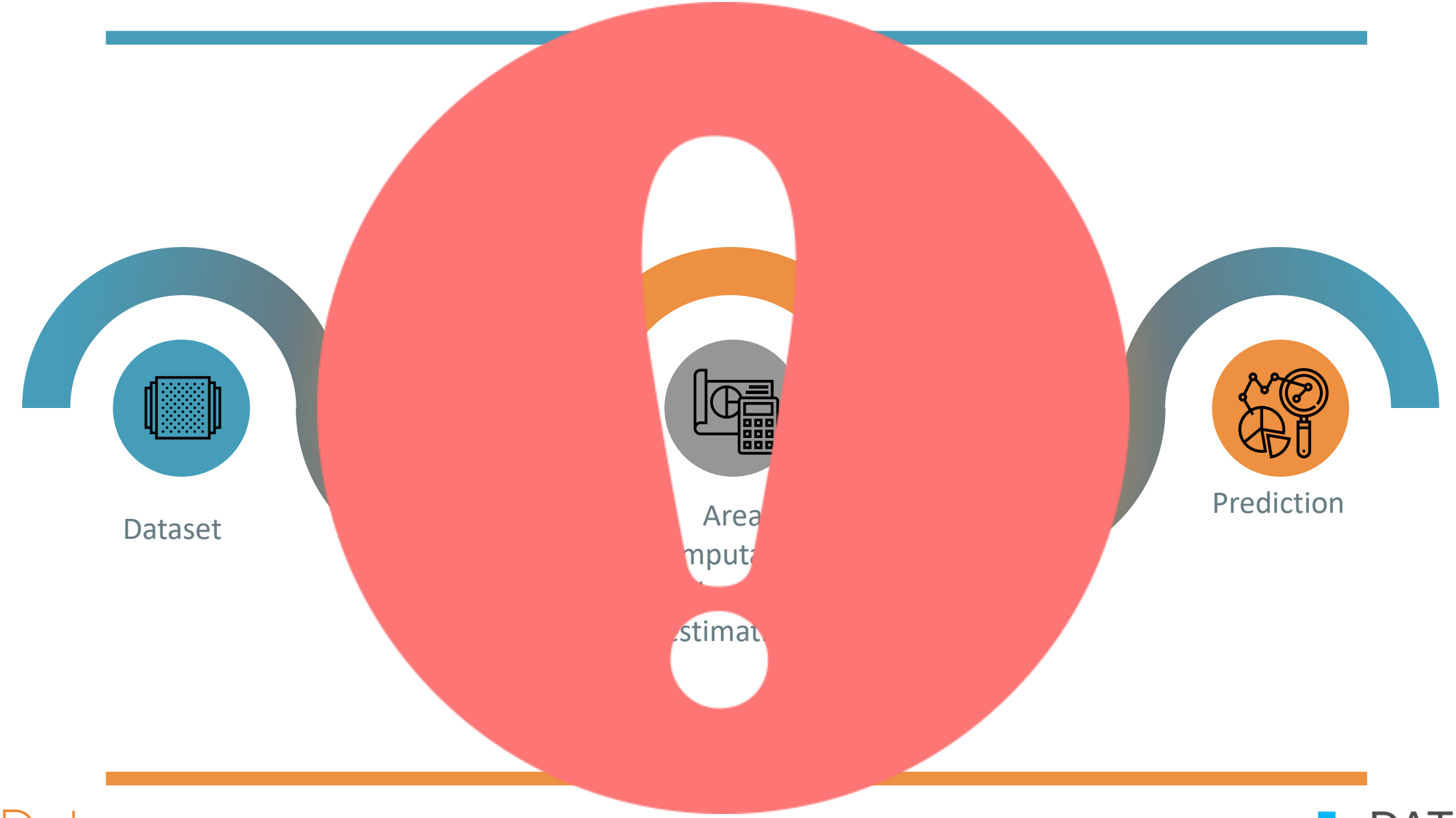
Is speeding up the medical research, and obtaining a correct diagnosis possible?

METHODOLOGY



PyData

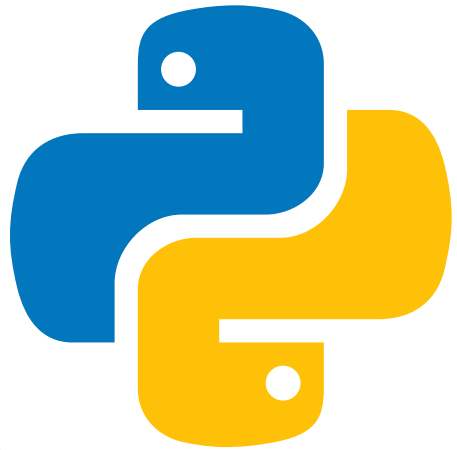
DATALAB
MEDIALAB USAL



PyData

DATALAB
MEDIALAB USAL

TECNOLOGÍAS EMPLEADAS



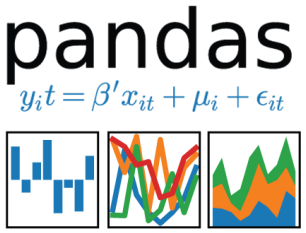
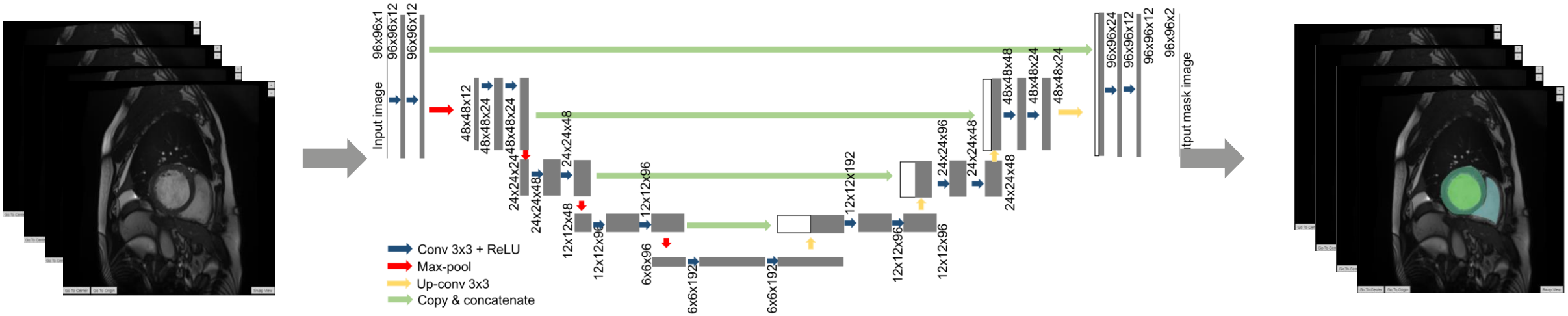


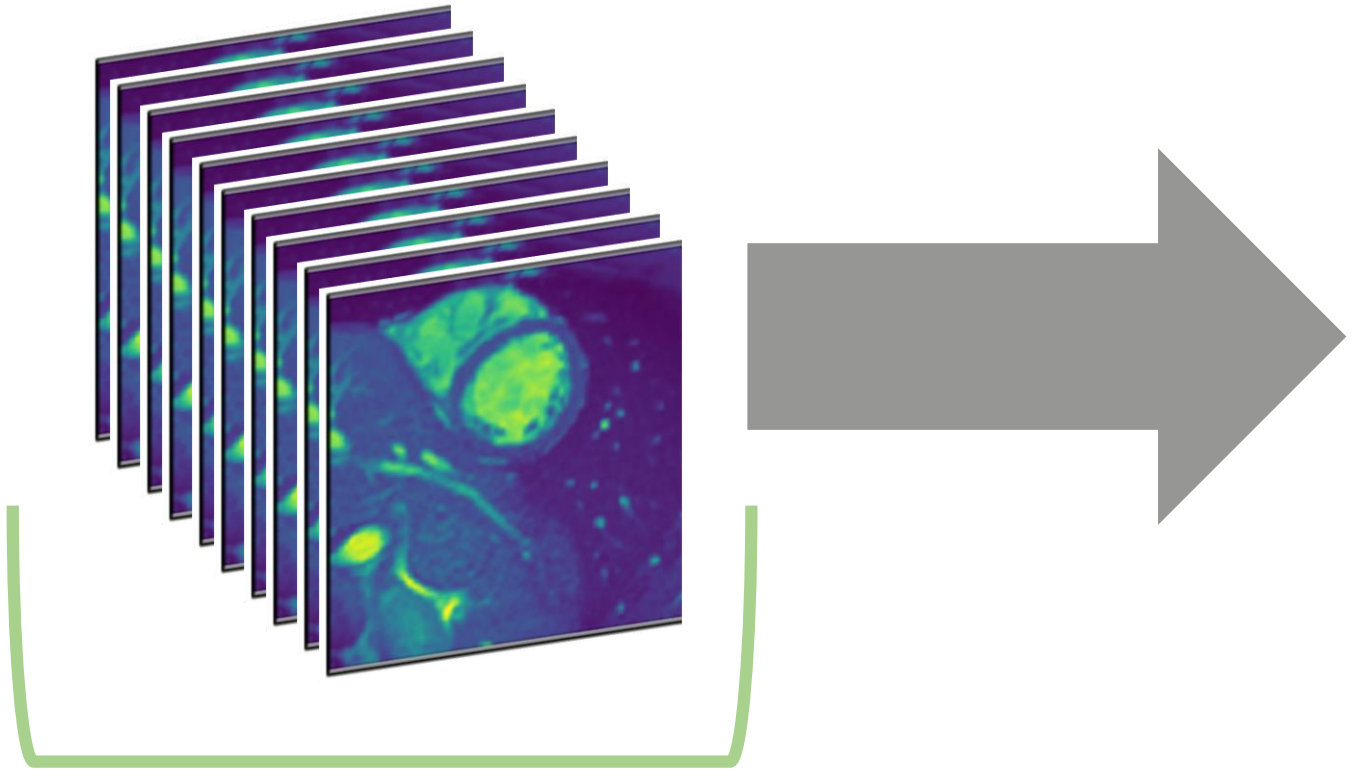
Image dataset and U-net



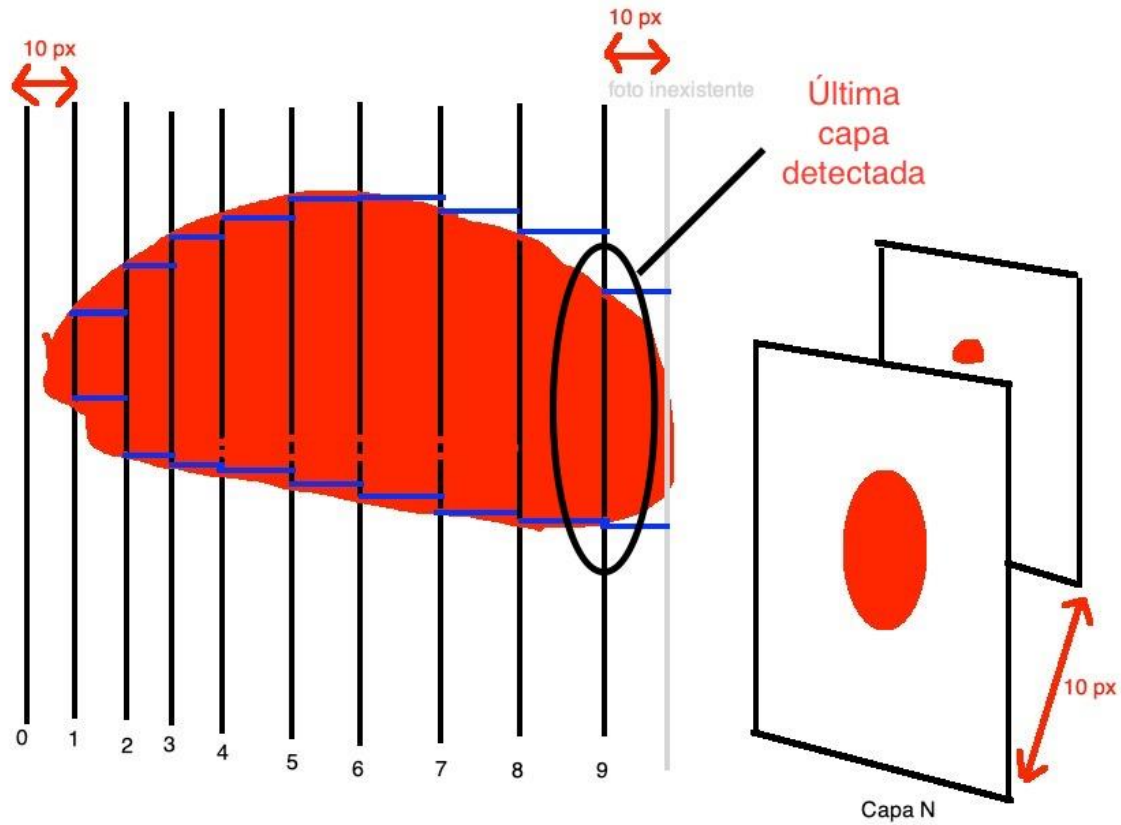
PyData

DATALAB
MEDIALAB USAL

Areas and volumes



- Area estimate for each slice
- Final volume estimate



VOLUMEN CALCULADO POR NOSOTROS

VOLUMEN REAL DEL MIOCARDIO (POR EJEMPLO)

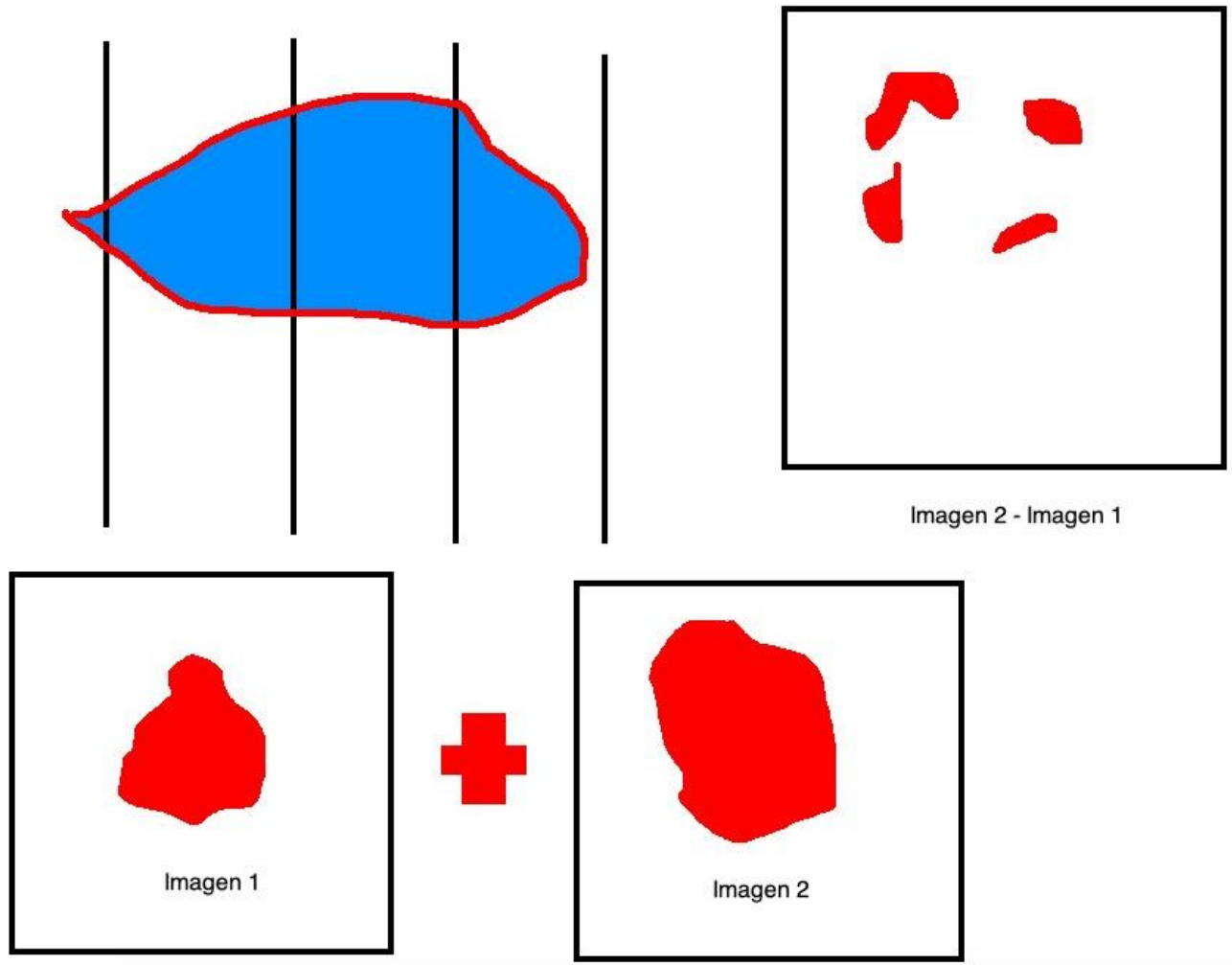
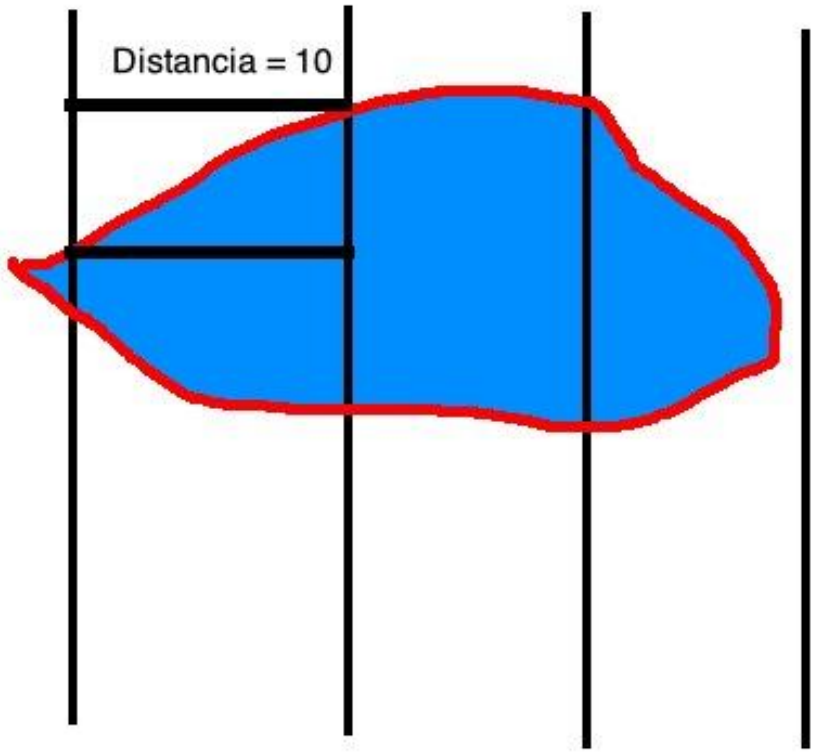


Imagen 2 - Imagen 1
Imagen 1
Imagen 2

Volumen cilindro/prisma = Area Base * distancia

Aproximadamente:
Volumen buscado = Volumen prisma / 2



Área de la base:

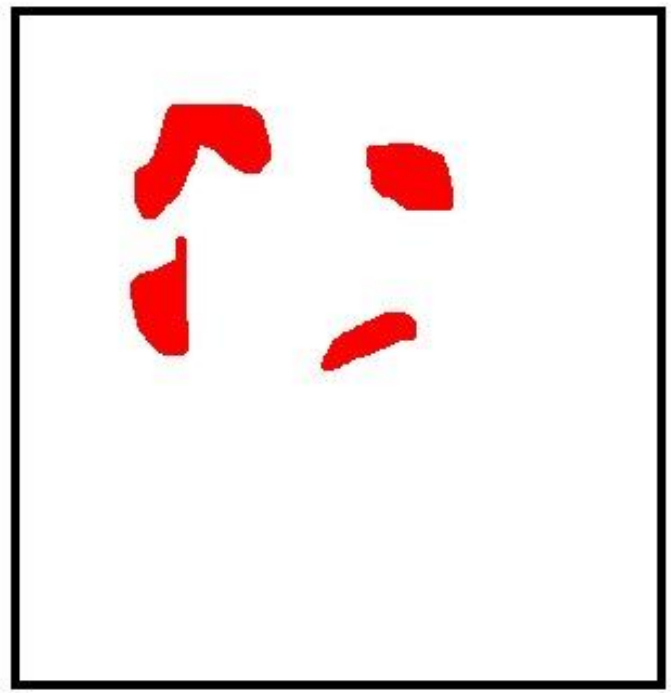


Imagen 2 - Imagen 1

```
In [421]: rf = RandomForestClassifier(n_estimators=10, max_leaf_nodes=5)

In [422]: rf.fit(train_X, train_y)

Out[422]: RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
                                 max_depth=None, max_features='auto', max_leaf_nodes=5,
                                 min_impurity_decrease=0.0, min_impurity_split=None,
                                 min_samples_leaf=1, min_samples_split=2,
                                 min_weight_fraction_leaf=0.0, n_estimators=10, n_jobs=1,
                                 oob_score=False, random_state=None, verbose=0,
                                 warm_start=False)

In [432]: y_pred = rf.predict(test_X)

In [433]: accuracy = accuracy_score(test_y, y_pred)
          accuracy

Out[433]: 0.9166666666666666

In [436]: log = log_loss(test_y, y_pred)
          log

Out[436]: 2.8782979993617404

In [434]: cv_results = cross_validate(rf, all_X, all_y, scoring='log_loss',
          ...                         return_train_score=False)
          cv_results

Out[434]: {'fit_time': array([0.02494311, 0.01840091, 0.01413107]),
           'score_time': array([0.00368381, 0.00176215, 0.00174999]),
           'test_score': array([-0.54772239, -0.49286839, -0.5179822 ])}
```

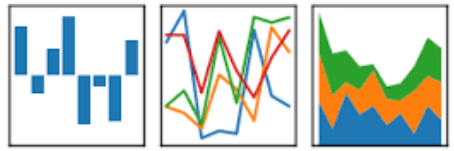
PyData

DATALAB
MEDIALAB USAL

Dataframe

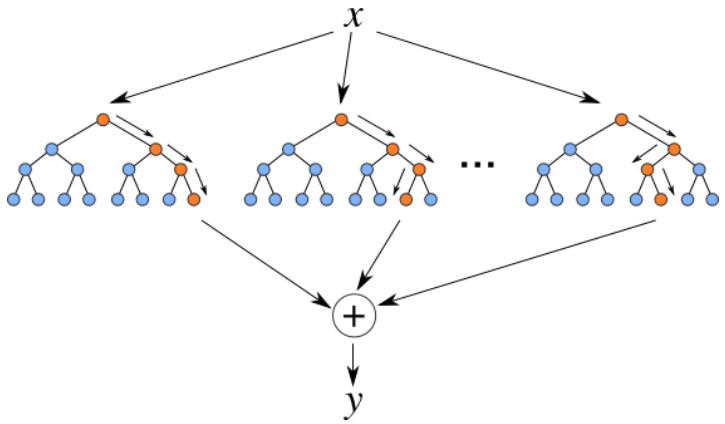
pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



sys_volume_RV_mL	sys_volume_LV_mL	sys_volume_Myo_mL	dia_volume_RV_mL	dia_volume_LV_mL	dia_volume_Myo_mL	patology	ey_frac_LV	ey_frac_RV
67.089844	228.369141	196.777344	146.289062	297.241211	177.612305	DCM	0.231704	0.541389
33.028793	182.957458	204.584122	99.348068	255.314255	172.209549	DCM	0.283403	0.667545
177.368164	245.239258	200.683594	189.306641	275.952148	196.020508	DCM	0.111298	0.063064
87.852478	227.444458	182.209778	107.740784	264.192963	166.115952	DCM	0.139097	0.184594
87.980713	224.846191	233.745117	168.822510	288.107666	202.757080	DCM	0.219576	0.478857
181.006622	279.636383	203.408432	280.872345	335.409164	203.130341	DCM	0.166283	0.355556
232.031250	279.527344	197.015625	303.363281	315.914062	193.007812	DCM	0.115179	0.235137
95.288086	237.890625	196.313477	168.896484	283.203125	180.590820	DCM	0.160000	0.435820
67.945729	229.034118	214.846826	86.675198	262.418182	204.603563	DCM	0.127217	0.216088
209.667969	266.723633	168.676758	282.031250	302.856445	173.046875	DCM	0.119307	0.256579

Predictions



0,01 %	29.49 %	7 %	63 %	0,5 %
RM	DCM		HCM	MINF
Abnormal right ventricle	Dilated cardiomyopathy		Left Ventricular hypertrophy	Myocardial infarction

Normal

NOR

PyData

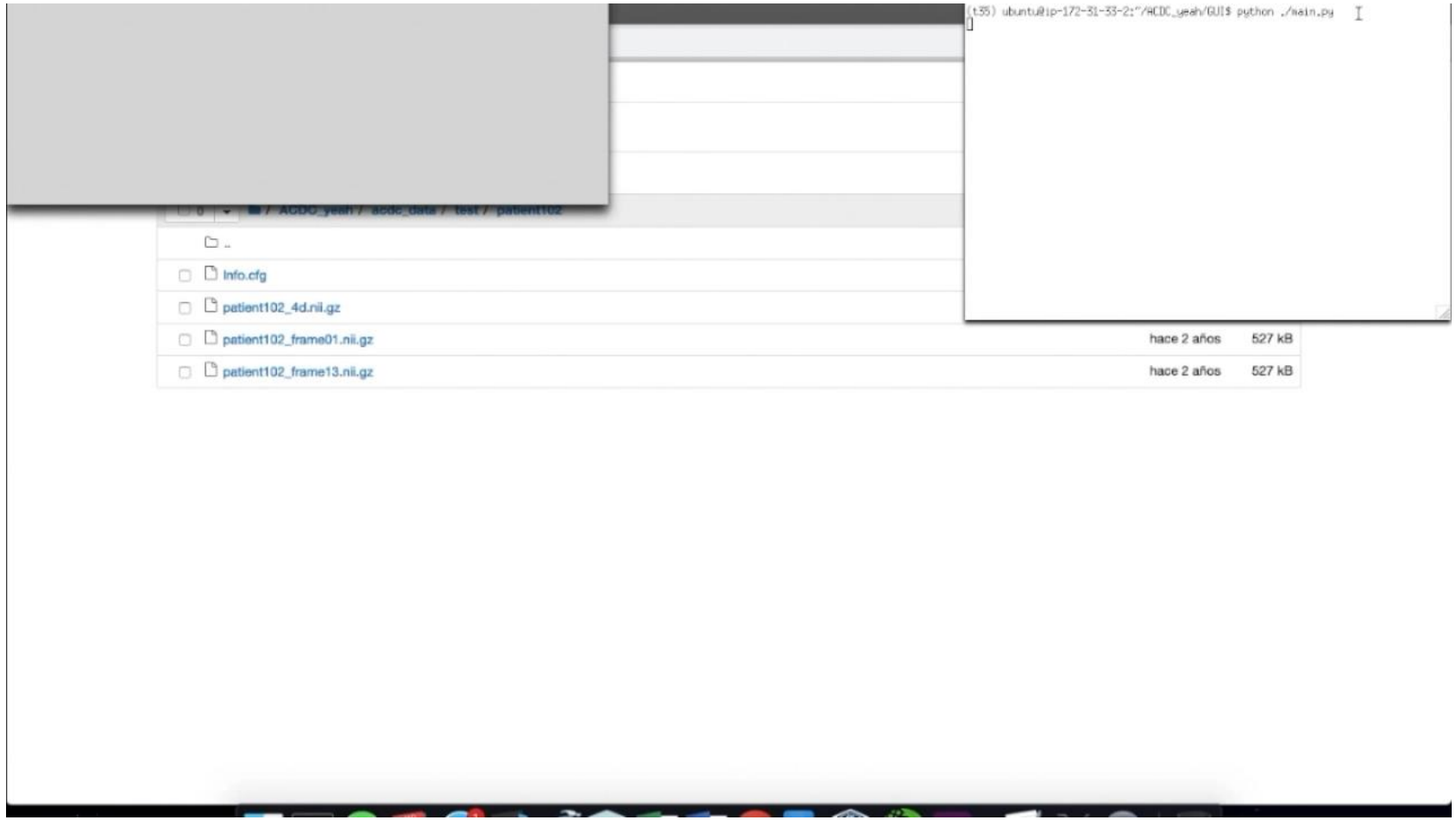
DATALAB
MEDIALAB USAL

AI USAGE ADVANTAGES

INSTANT DIAGNOSIS

HIGH ACCURACY (82%)

MORE TIME IS SPENT ON EACH PATIENT



PyData

DATALAB
MEDIALAB USAL

THANK YOU FOR YOUR ATTENTION

ANY QUESTIONS?