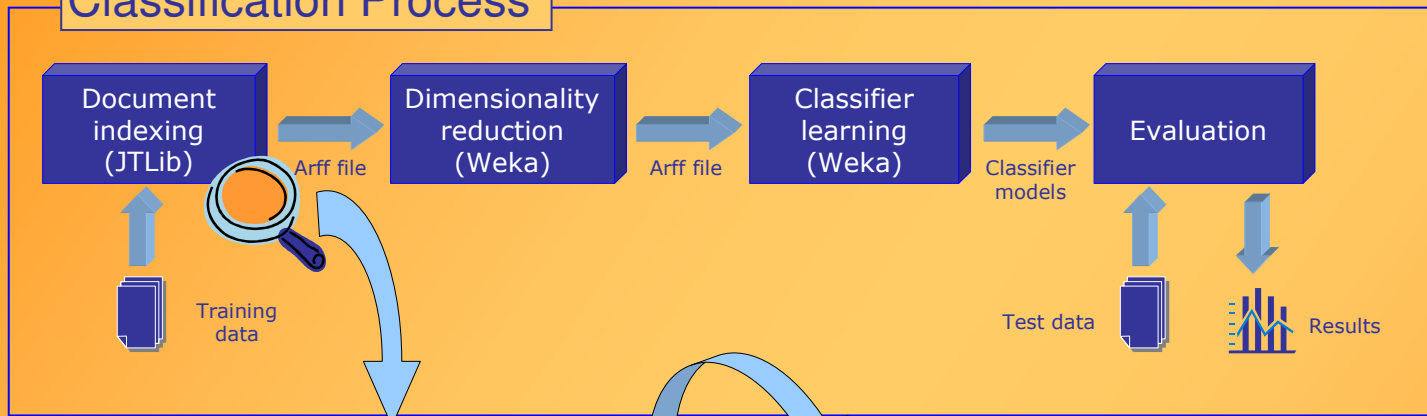


# Feature Engineering and Quick Prototyping of PPI Classifiers\*

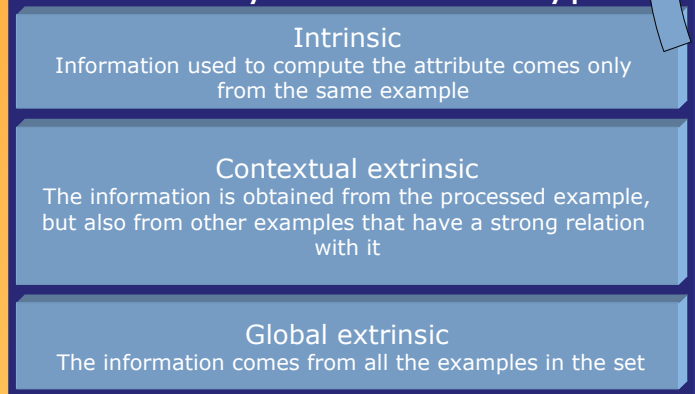
Manuel J. Maña López  
Jacinto Mata Vázquez  
Universidad de Huelva  
{manuel.mana, mata}@diesia.uhu.es

Francisco Carrero García  
José M<sup>a</sup> Gómez Hidalgo  
Universidad Europea de Madrid  
{francisco.carrero, jmgomez}@uem.es

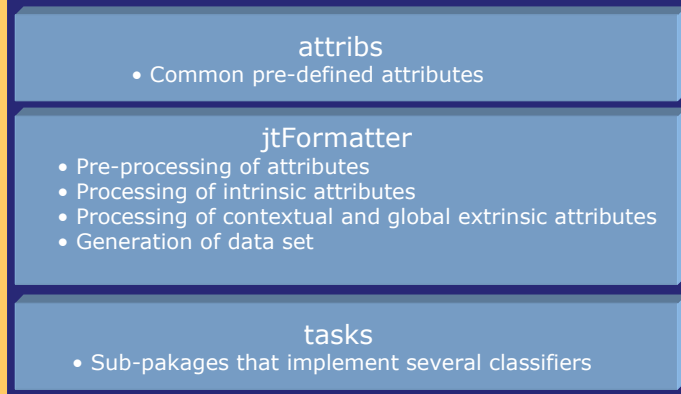
## Classification Process



## Taxonomy of Attribute Types



## JTLib



## Selected Attributes

Attribute type	Attribute Name
<b>Global extrinsic</b>	Unigrams Bigrams Trigrams

## Ranking of Attributes

N-tuples with higher correlation coefficient		
interaction	two hybrid	yeast two hybrid
domain	domain of	two hybrid system
binding	yeast two	we show that
hybrid	with the	the yeast two
yeast	interacts with	a yeast two
with	in vitro	the interaction of
interacts	the interaction	to interact with
terminal	interact with	in vitro and
complex	interaction between	is required for
interact	interaction with	two hybrid screen

N-tuples with lower correlation coefficient		
ms	<number> de	dimensional gel electrophoresis
de	two dimensional	two dimensional gel
spots	protein spots	by <number> de
proteome	<number> d	malDI TOF MS
proteomic	gel electrophoresis	<number> protein spots

## Best Run Results

	Adaboost with Naïve Bayes N° of unigrams/bigrams/trigrams 5140 / 18510 / 10224
<b>Precision</b>	0.583
<b>Recall</b>	0.952
<b>F-measure</b>	0.723

\* Funding: Spanish Ministry of Education and Science and the European Union from the European Regional Development Fund (ERDF) - (TIN2005-08988-C02-01 and TIN2005-08988-C02-02); Madrid Regional Government under the IV PRICIT programme through the Joint Research Programme MAVIR (S-0505/TIC-0267).