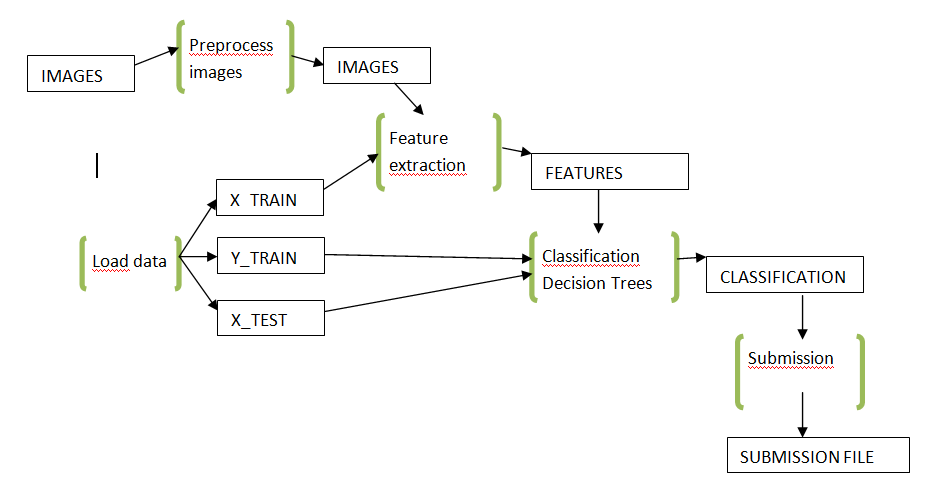
**Process description**



**Process:** load data Diede

Input: csv files

Output: dictionary ‘data’ with fields: X\_TRAIN, Y\_TRAIN and X\_TEST

**Process:** preprocess images Roos

Input: raw .jpg images

Output: clean .jpg images

**Process:** feature extraction Daniëlle: input. Laurens: features.

Input: images, X\_train

Output: features

**Process:** classification Diede

Input: features, Y\_train, X\_test

Output: classification

**Process:** submission Diede

Input: classification

Output: submission file

**Files:**

**X\_TRAIN:**

business\_id Photo\_id

1000 3

1000 505

1000 9809

1001 57

… …

**Y\_TRAIN:**

business\_id C1 C2 C3 …

1000 1 0 1

1001 1 1 0

1002 0 0 0

… … … …

**X\_TEST:**

business\_id Photo\_id

2000 4

2000 506

2000 9810

2001 58

… …

**FEATURES:**

business\_id F1 F2 F3 …

1000 0.89 0 0.65

1001 0.1 1 0.98

1002 0.23 0 0.23

… … … …

**CLASSIFICATION:**

business\_id L1 L2 L3 …

1000 1 1 1

1001 0 1 0

1002 0 0 1

… … … …

**SUBMISSION FILE:**

nog uitzoeken