

Projet Apprentissage Approfondi

2SN - Logiciel

---

# Rapport sur la constitution de Base de données

---

Réalisé par :

- Qingwei Jia
- ECH-CHEBLAOUI Yassine
- AMGHAR Souhail

Encadrant :

- Pr. Carlier Axel

## Table de matière

1. Introduction.....	3
2. Description du Sujet.....	3
3. Description de méthodologie d'acquisition des données.....	4
4. Description de méthodologie de répartition des données.....	4
5. Pronostic.....	5
6. Script de chargement des données .....	5
7. Base de Données .....	5

# 1. Introduction

Notre Groupe est composé de 3 personnes de L-SN :

- Ech-cheblaoui Yassine
- Qingwei Jia
- Amghar Souhail

On travaille sur le sujet de :

« Reconnaissance des billets et pièces en euros entre 1 centime – 100 € »

Lien Git :

<https://github.com/Testoxe/EuroMoneyRecognition/>

## 2. Description du Sujet

Depuis la création de la monnaie unique européenne, il existe huit billets en euros répartis sur deux séries, dont les valeurs faciales s'échelonnent de 5 à 500 euros et huit pièces de monnaies (1 cent, 2 cent, 5, 10, 20, 50, 1€, 2€). *(ce qui font les champs de notre classification de la DB par la suite).*

De l'introduction de ces billets le 1er janvier 2002 au 1er mai 2013, une seule série a été en circulation, mais une deuxième série a commencé à remplacer celle-ci à partir du 2 mai 2013.

**Clarification :** Notre travail fait l'interpellation et l'identification de la nouvelle série de monnaie (post 2013).

Contrairement aux pièces en euros, le design des billets est le même pour l'ensemble de la zone euro. Ils sont cependant imprimés et émis par les différents États membres. Quel que soit le pays émetteur, les billets sont cependant utilisables dans tous les pays dont l'euro est la monnaie. Ils sont conçus en pure fibre de coton, ce qui leur donne une meilleure résistance et ce qui permet de les reconnaître facilement au toucher. Jusqu'à présent, la taille des billets s'échelonne d'un format de 120 sur 62 millimètres pour le billet de 5 euros

à 160 sur 82 millimètres pour le billet de 500 euros. Ils ont tous une couleur différente, en particulier avec les billets de valeurs proches, afin de pouvoir bien les distinguer. Ce qui fait effectivement une sorte de bruit (eng. noise) naturel le donnée d'apprentissage.

**Clarification :** Pour les pièces de moins de 5€, on ferait l'entraînement de la base de donnée sur celles de la France.

### 3. Description de méthodologie d'acquisition des données

Le regroupement de la base de données, c'est fait par l'acquisition et regroupement des images gratuits, libre de droit retrouvées dans le moteur de recherche Google Image.

Après l'identification des images adéquates, on a effectué deux types de script :

- Script en JavaScript pour identification des images, regroupement de la sélection et téléchargement.
- Script en Python pour classification des ressources.

### 4. Description de méthodologie de répartition des données

Ainsi téléchargé, la base de données est répartition sur 3 ensembles d'entraînement, de validation et de test.

L'ensemble d'apprentissage lui est réparti dans des fichiers suivant la catégorie des monnaies :

- On a créé des fichiers selon les groupements de monnaies depuis 1cent jusqu'à 100€.
- Pour la nomination des images dans les fichiers, ils suivent une syntaxe incrémentale simple.

## 5. Pronostic

Nous estimons que le défi que nous devons relever avec notre projet d'apprentissage approfondi n'est pas de simple.

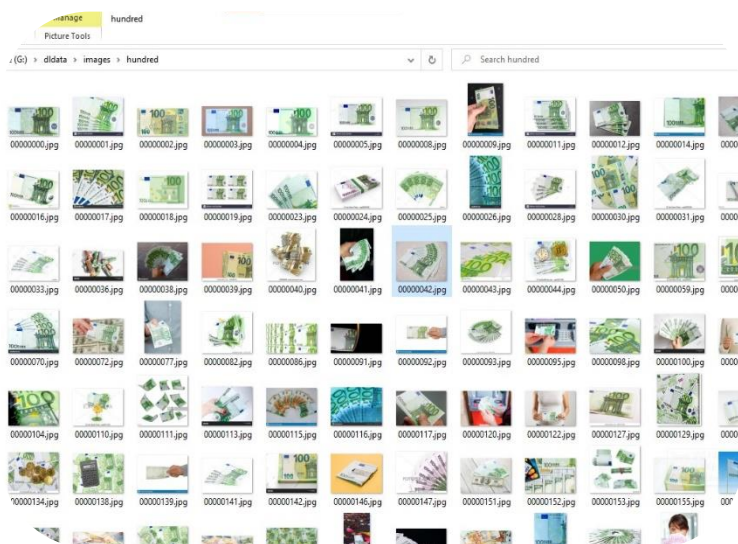
Le projet répond à un apprentissage supervisé de type **Classification multinomiale** avec un grand problème de bruit sur la base de données. Surtout avec des catégories de billets qui se ressemblent (même couleur, même taille, même font et même design) qui rend l'apprentissage plus challengeant et complexe.

## 6. Script de chargement des données

(Voir le git)

## 7. Base de Données

(Voir le git)



nomination

	Name	Date modified	Type	Size
access				
ktop	5	2021-03-22 17:53	File folder	
vnloads	10	2021-03-22 17:53	File folder	
	20	2021-03-22 17:53	File folder	
ures	50	2021-03-22 17:53	File folder	
.N7	100	2021-03-22 17:53	File folder	
	cinq	2021-03-22 17:53	File folder	
	fifty	2021-03-22 17:53	File folder	
glais	five	2021-03-22 17:53	File folder	
	hundred	2021-03-22 17:53	File folder	
	ten	2021-03-22 17:53	File folder	
	twenty	2021-03-22 17:53	File folder	
	vingt	2021-03-22 17:53	File folder	
ii-Projet				
ive Cloud Files				
rive - Ministere de				
rive - Personal				

classification

