**Power BI est une suite d’outils d’analyse commerciale permettant d’analyser des données et de partager des informations.** Les tableaux de bord Power BI offrent aux utilisateurs professionnels une vue globale avec leurs mesures les plus importantes au même endroit, mises à jour en temps réel et disponibles sur l’ensemble de leurs appareils. En un clic, les utilisateurs peuvent explorer les données sur lesquelles s’appuie leur tableau de bord en utilisant des outils intuitifs qui facilitent la recherche de réponses. Il est facile de créer un tableau de bord grâce à [des centaines de connexions](https://powerbi.microsoft.com/fr-fr/#connect-wrapper) à des applications de gestion des informations professionnelles populaires, complétées par des tableaux de bord prédéfinis qui vous permettent d’être rapidement opérationnel. Et vous pouvez **accéder à vos données et rapports où que vous soyez, à l’aide des applications**[**Power BI Mobile**](https://powerbi.microsoft.com/fr-fr/mobile/)**, qui se mettent à jour automatiquement avec toutes les modifications apportées à vos données.**

Si vous êtes analyste de données et que vous produisez des rapports et des analyses pour votre organisation, Power BI vous permet d’être productif et de faire preuve de créativité dans ce que vous créez. [**Power BI Desktop**](https://powerbi.microsoft.com/fr-fr/desktop/)**est un outil de création d’applications web hybride de données et de rapports riches en fonctionnalités .** Combinez des données provenant de plusieurs bases de données, fichiers et services web à l’aide d’outils visuels qui vous permettent de comprendre et de corriger automatiquement la qualité des données et les problèmes de mise en forme. Avec plus de 20 éléments visuels intégrés et un ensemble riche de visualisations personnalisées, créez d’incroyables rapports qui faire passer efficacement votre message. **Utilisez le**[**service Power BI**](https://powerbi.microsoft.com/fr-fr/features/)**pour publier des rapports en toute sécurité à votre organisation et configurer une actualisation automatique des données afin que tout le monde dispose des informations les plus récentes.**

Power BI peut unifier toutes les données de votre organisation, qu’elles se trouvent dans le cloud ou qu’elles soient stockées localement. Les [passerelles Power BI](https://powerbi.microsoft.com/fr-fr/gateway/) vous permettent de connecter des bases de données SQL Server, des modèles Analysis Services, ainsi que de nombreuses autres sources de données à vos tableaux de bord dans Power BI. Si vous disposez déjà de portails ou d’applications de rapport, [incorporez](https://powerbi.microsoft.com/fr-fr/developers/) les rapports et des tableaux de bord Power BI pour une expérience unifiée.

<https://www.societegenerale.com/rapport-d-activite/rapport-activite-fr.html>

<https://www.societegenerale.com/sites/default/files/documents/Document%20de%20r%C3%A9f%C3%A9rence/2017/Societe-Generale-DDR-2017-depot-amf-08032017-FR.pdf>

2. PRÉSENTATION DE SOCIÉTÉ GÉNÉRALE Société Générale est l’un des tout premiers groupes européens de services financiers. S’appuyant sur un modèle équilibré et diversifié, le Groupe allie solidité financière et stratégie de croissance durable, afin de mettre sa performance au service du financement de l’économie et des projets de ses clients. Avec un ancrage solide en Europe et une présence dans les géographies à fort potentiel, les 145 700 collaborateurs\* du Groupe, présents dans 67 pays, accompagnent au quotidien 31 millions de clients particuliers, grandes entreprises, investisseurs institutionnels(1) , dans le monde entier, en offrant une large palette de conseils et de solutions financières qui s’appuie sur trois pôles métiers complémentaires :

■ la Banque de détail en France avec les enseignes Société Générale, Crédit du Nord, et Boursorama qui proposent des gammes complètes de services financiers avec une offre multicanal à la pointe de l’innovation digitale ;

■ la Banque de détail à l’International, l’Assurance, et les Services financiers aux entreprises, avec des réseaux présents dans les zones géographiques en développement et des métiers spécialisés leaders dans leurs marchés ;

■ la Banque de Financement et d’Investissement, Banque Privée, Gestion d’Actifs et Métier Titres, avec leurs expertises reconnues, leurs positions internationales clés et des solutions intégrées. Au 7 mars 2017, la notation long terme de Société Générale est A (FitchRatings, Standard and Poor’s et R&I), A (High) (DBRS), A2 (Moody’s). Société Générale figure dans les principaux indices de développement durable : DJSI (World et Europe), FTSE4Good (Global et Europe), Euronext Vigeo (Europe, Eurozone et France), ESI Excellence (Europe) d’Ethibel, quatre des indices STOXX ESG Leader et MSCI Low Carbon Leaders Index.

**Goal :**

Prediction of the GL account to be used on the non PO based invoice lines at the APC/AP1 step

**Scope:**

-Non PO based invoices

-SG Metro without GBIS (to start with)

A l’heure où les voitures deviennent autonomes, où les géants du web se sont lancés dans une course folle à l’élaboration d’assistants virtuels, les Centres de Services Partagés DFIN/DOM/PTP ne peuvent-ils pas tirés profit des nouvelles technologies ?

Sachant que les CSP DOM paient plus de 6 Milliards d’euro de factures fournisseurs par an et que le périmètre de ces CSP augmente d’année en année.

Sachant que pour certaines entités nous disposons de plus de 15 ans d’historique.

Ce volume de données conséquent n’est-il pas une formidable source pour développer des algorithmes de machine learning qui permettront d’optimiser et d’automatiser le traitement des nouvelles factures ?

Un exemple concret est la prédiction des informations à compléter sur les nouvelles factures : compte comptable, TVA…, en fonction des données historiques.

A long terme, ces algorithmes, couplés à nos initiatives de déploiement de robot (RPA) ne permettraient-ils pas de basculer les activités Procure to Pay dans l’ère de l’intelligence artificielle et du Smart Process Automation ?

Nous pensons que oui, c’est l’ambition et l’initiative que nous mettons en avant via cette idée.

Difficulté :

445 compte différente, et des nouveaux comptes à ajouter au fur et au mesure

Si on regroupe les données par fournisseur directement, certains fournisseurs ont très peu de données à apprendre

Les « feature » ne sont pas fortement liée avec le compte

Valeur de TVA, Business\_nature etc sont des valeurs par défaut

Stratégie :

Plan du compte :

Expense :

FRA01 : compte PCI

SHARE : compte E2K (hors Lyxor) Alt\_Compte plus commence par non 9

996- 6xxxxxxx

993- 3xxxxxxxx

SHPCG : compte Lyxor (S4892 et S4893)

Expense Bis :

Compte PCI

996xxxxx :

993xxxxx : compte de charge

PCI commence par 9 BU est filial

Cas particulier :

1. Nouveaux fournisseurs
2. Fournisseur a un seulement compte dans les données historiques
3. Lissage

Features :

|  |  |  |
| --- | --- | --- |
| Feature | Type | Description |
| BUSINESS\_UNIT | Catégoriel | Code entité |
| FILIAL |  | Filale |
| VENDOR\_SETID |  |  |
| VENDOR\_ID |  |  |
| MONETARY\_AMOUNT | Continu |  |
| VAT\_INV\_AMT\_GL |  |  |
| ASST\_FLG |  |  |
| VAT\_TREATMENT |  |  |
| VAT\_SERVICE\_TYPE |  |  |
| VOUCHER\_ID |  |  |
| TAX\_CD\_VAT |  | Code tva |
| DESCR |  | Description de prestation |
| ACTIVE\_FLG |  | Marque de lissage |
| VNDR\_FIELD\_C30\_J |  |  |
| PROB\_VENDOR\_C1 |  |  |
| PROB\_VENDOR\_C3 |  |  |
| PROB\_VENDOR\_C993 |  |  |
| PROB\_VENDOR\_C997 |  |  |
| VENDOR\_ONLY |  |  |
| VENDOR\_ONLY\_993 |  |  |
| VENDOR\_ONLY\_996 |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Modèle de classification :

1. Modèle de classification du premier chiffre du compte
2. Modèle de classification de 3 premiers chiffres du compte (993, 996, 994 etc)
3. Modèle de classification pour les comptes commencent par 996 :

Traitement par fournisseur

1. Modèle de classification pour les comptes commencent par 993/994/997

Implémentation :

1. Langage de programmation : Python 3.5
2. Algorithme de machine Learning : Random Forest
3. Application : GUI ( Graphical User Interface) et générer un exécutable avec cx\_freeze

Pour distribuer l’outil

Random Forest :

Random Forest est un algorithme de machine learning qui est particulièrement efficace pour repérer des liens entre une variable à expliquer et des variables explicatives. Random Forest va classer les variables explicatives en fonction de leurs liens avec la variable à expliquer.

GUI tkinter :

Tkinter est une boite à outils pour faire des interfaces graphiques en Python. Elle repose sur un autre langage de script, le Tcl (Tool Command Language) en utilisant la bibliothèque Tk de ce langage. l'un des avantages de Tkinter est sa portabilité sur les OS les plus utilisés par le grand public.

Cx\_Freeze :

Cx\_Freeze permet de retrouver avec un fichier .exe et plusieurs fichiers compagnons de fichier .py, bien plus faciles à distribuer et, pour les utilisateurs, à exécuter. De cette façon, on peut transmettre le programme sans obliger les utilisateurs à installer Python sur leur ordinateur.

Avantages de cx\_Freeze

* Portabilité : cx\_Freeze est fait pour fonctionner aussi bien sur Windows que sur Linux ou Mac OS ;
* Compatibilité : cx\_Freeze fonctionne sur des projets Python de la branche 2.X ou 3.X ;
* Simplicité : créer son programme standalone avec cx\_Freeze est simple et rapide ;
* Souplesse : nous pouvons aisément personnaliser le programme standalone avant de le construire.

Evaluation :

1. Matrix de Confusion
2. Temps de calcule

Prédiction du 3 premier chiffre : 2454 lignes 0.444s

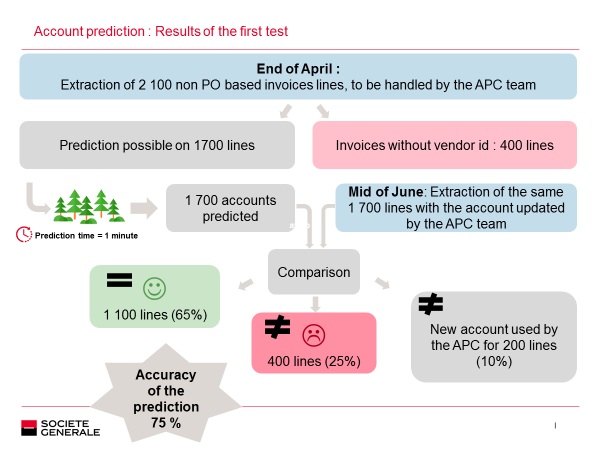
Prédiction du compte 993XXXXX : 0.299 s

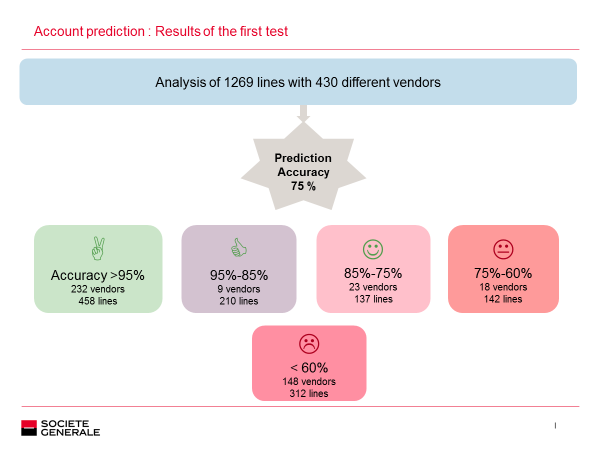
Prédiction du compte 996XXX : 534 fournisseur 120.139 s

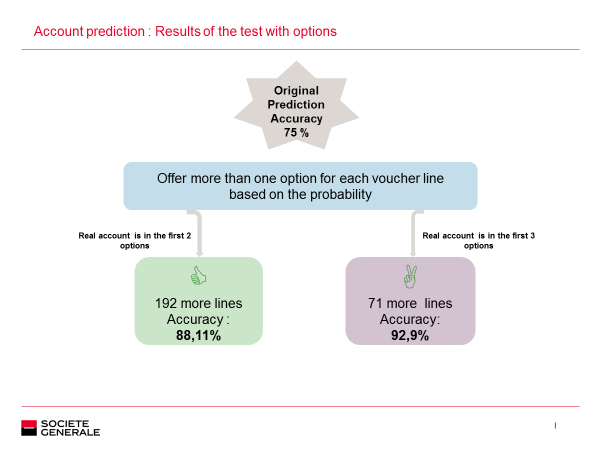
1. Test

Facture\_récyclée d’avril

|  |
| --- |
| Predict the first number : 9 or not  if c1=9:  Filter new vendor  Predict the firt 3 numbers 996,993 or others  filter vendor if only have one account in the historic  if c3=993:  filter vendor if only have one account begin with 993 in the historic  Predict 993XXXX  else if c3=996:  filter vendor if only have one account begin with 996 in the historic  for each Vendor:  predict 996XXXX  Offer the probability of each option(possible account) |







|  |  |  |  |
| --- | --- | --- | --- |
| **Title** | **Précision** | **Erreur** | **Description** |
| **Predicted\_1 :**  **9 or not** | 1834/ 1837    99.84% | 3 lines of Z0010(filial) use the  account 996XXX  Index (968,969,970) | Filtre :  Lignes en attant |
| **Predicted\_3 :**  **993 ; 996; autre** | 1437/1480    97.09% | Between 993 et 996:  35 lines  Prediction is 993 but real is 996 : 31 lines | Filtre:  NEW\_VENDOR  FILIAL  NEW\_ACCOUNT |
| **Predcited\_all**    **993XXX**  **996XXX(par fournisseur)** | 1112/1480  75.14% |  |  |
| 1109/1437  77.17% |  | Filtre :  Wrong first three numbers |
| 168  100% |  | VENDOR\_ONLY  (168 lignes) |
| 941/1269  74.15% |  | Filtrer VENDOR\_ONLY |

Résultat Expese AEX:

Données : facture comptabilisé du mai

Nombre de ligne total : 4714

Nombre de ligne avec « New Vendor » : 277

Ligne avec « ALTACCT » : 4247

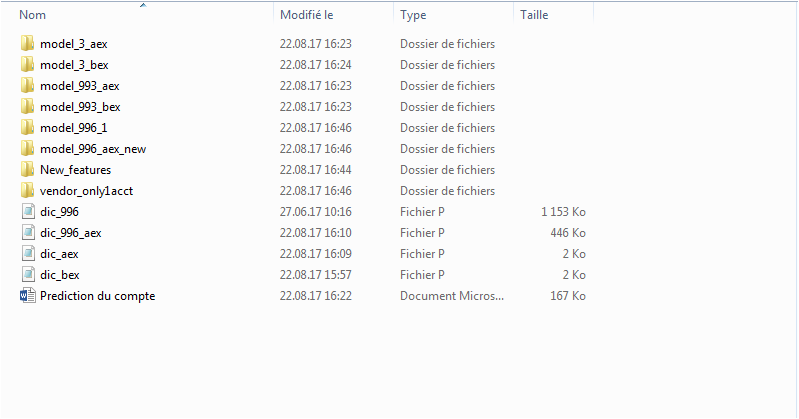
Ligne de prédiction à vérifier : 4090

Précision :

Option 1 ou valeur par défaut : 81 % (3314 sur 4090)

Option 2 : 82% (3314 + 57 sur 4090)

**Annexe :**



* Models : modèle de Forêt aléatoire
* Dics : Dictionnaire contiens les listes des colonnes pour chaque modèle
* New\_Feature : fichier excel , probabilité de chaque compte par fournisseur
* Vendor\_ony1acct : fichier excel de fournisseur qui a un seul compte dans les données historique

1. Nouveaux comptes

Perspective :

1. System recommandation

Société Générale est l’un des tout premiers groupes européens de services financiers. S’appuyant sur un modèle équilibré et diversifié, le Groupe allie solidité financière et stratégie de croissance durable, afin de mettre sa performance au service du financement de l’économie et des projets de ses clients. Avec un ancrage solide en Europe et une présence dans les géographies à fort potentiel, les 145 700 collaborateurs\* du Groupe, présents dans 67 pays, accompagnent au quotidien 31 millions de clients particuliers, grandes entreprises, investisseurs institutionnels (1) , dans le monde entier, en offrant une large palette de conseils et de solutions financières qui s’appuie sur trois pôles métiers complémentaires :

■ la Banque de détail en France avec les enseignes Société Générale, Crédit du Nord, et Boursorama qui proposent des gammes complètes de services financiers avec une offre multicanal à la pointe de l’innovation digitale ;

■ la Banque de détail à l’International, l’Assurance, et les Services financiers aux entreprises, avec des réseaux présents dans les zones géographiques en développement et des métiers spécialisés leaders dans leurs marchés ;

■ la Banque de Financement et d’Investissement, Banque Privée, Gestion d’Actifs et Métier Titres, avec leurs expertises reconnues, leurs positions internationales clés et des solutions intégrées. Au 7 mars 2017, la notation long terme de Société Générale est A (FitchRatings, Standard and Poor’s et R&I), A (High) (DBRS), A2 (Moody’s). Société Générale figure dans les principaux indices de développement durable : DJSI (World et Europe), FTSE4Good (Global et Europe), Euronext Vigeo (Europe, Eurozone et France), ESI Excellence (Europe) d’Ethibel, quatre des indices STOXX ESG Leader et MSCI Low Carbon Leaders Index.

Plan :

Data vis :

La data visualisation (Data Viz ou visualisation des données) est une technique qui permet de présenter des milliers de chiffre et de données de façon visuelle, interactive et dynamique afin de les rendre facilement compréhensible et d’en explorer toutes les dimensions.

Lors que la présentation, nous allons présenter trois outils de type différentes qui sont tous disponibles dans notre PC, accompagnant une démo (visualisation des données de factures Paris-Bucharest ) pour illustrer l’utilisation de dataviz dans le cas concrète chez PTP, surtout dans votre équipe.