

Livret des stages

Table of contents

1	Livret des stages	5
2	Présentation du master SEP	7
3	Administration publique	9
3.1	Conseil Regional Nouvelle Aquitaine	9
3.2	Union départementale des centres communaux d'action sociale de la Seine-Saint-Denis (UDCCAS)	10
4	Aéronotique, Espace, défense et sécurité	11
4.1	Groupe Safran	11
5	Banque et Assurance	13
5.2	AXA France	14
5.3	BNP Paribas Cardif	15
5.4	Generali	15
5.5	GROUPAMA	16
5.6	MGP	17
5.7	BANQUE POPULAIRE VAL DE FANCE	17
5.9	BNP PARIBAS personal Finance	18
5.10	BNP PARIBAS Securites Services	19
5.11	Caisse d'épargne Bretagne pays de loire	19
5.13	Credit Agricole du Nord-Est	21
5.14	LE CREDIT LYONNAIS (LCL)	23
5.15	Oney Bank	23
5.16	RCI BANK AND SERVICES	24
5.17	RCI banque -DIAC	25
5.18	Société Générale	26
6	Commerce et Industrie	28
6.1	Arcelor Mittal	28
6.2	BLT Cosmetics LABOTÉ	29
6.3	Chambre de commerce et d'industrie des Ardennes	30
6.4	PUM Plastiques	30
6.5	Zenpark	31

7	Conseil	33
7.1	BGFi Consulting	33
7.2	BI Consulting	34
7.3	Business & Decision	34
7.4	Capgemini	37
7.5	CERFRANCE	38
7.6	CGI	39
7.7	CROSS SYSTEMS	40
7.8	Data Company	40
7.9	Data Vectis	41
7.10	DSI-Consulting	42
7.11	Efficience 3	43
7.12	Equancy	44
7.13	Greenflex	44
7.14	Greenomy	45
7.15	Groupe AVISIA	46
7.16	Hanalytics	47
7.17	Innogen Consulting	48
7.18	JANASENCE	49
7.19	Kheopsys Lab	49
7.20	QUANT AI LAB	50
7.21	Umanis SA	51
7.22	Velvet Consulting	51
8	Eau et Assainissement	53
9	Economie et développement	54
9.1	Centre d'études de politiques pour le développement (CEPOD)	54
10	Edition de logiciels	56
11	Energie	61
11.1	Akuo Energy SAS	62
11.2	EDF	62
11.3	ENEDIS	63
11.4	ENGIE SA	63
11.5	Kayros	64
11.6	PETROSEN TRADING & SERVICES	65
11.7	VEOLIA	65
11.8	Voltalis	66
12	Grande distribution	68
12.1	Siège mondial Carrefour	68

13 Hôtellerie	70
14 Informatique et Télécommunication	71
14.1 Data moove	71
14.2 Line Up Data Science	72
14.3 Linkfluence FRANCE	73
14.4 Monstock	73
14.5 Mybestpro (telecommunications)	74
14.6 Orange SA	74
15 Luxe	77
15.1 Moët Hennessy	78
16 Média	79
16.1 Havas Média France	79
16.3 Phenix	81
16.4 Prisma média	82
17 Recherche	83
17.1 Centre de Calcul Régional ROMEO (Université de Reims Champagne Ardenne)	84
17.2 Centre de recherche en STIC (CReSTIC) / Binus university	85
17.4 Institut de Recherche bio-Médicale et d'Épidémiologie du Sport	87
17.5 Laboratoire de Mathématiques de Reims	87
17.6 Lab'URBA Université Paris-Est	88
17.7 NEOMA Business School	89
17.8 Université de Paris, MAP5 (CNRS UMR 8145)	89
17.9 UPEM (Université Paris-Est-Marne-la-vallée)	90
18 Services de la santé	91
18.1 Cosmetic Dermatology LLC (DR BRANT)	91
18.2 KAP CODE	92
18.3 Logever Healthcare	93
19 Voyage & Tourisme	94
20 Autres	95
20.1 CENTRE VINICOLE CHAMPAGNE NICOLAS FEUILLATTE (CV-CNF)	95
20.2 Daimler AG	96
20.3 Pari Mutuel Urbain (PMU)	97

1 Livret des stages



LIVRET DES STAGES

2022 / 2023

MASTER 2 : STATISTIQUE POUR
L'ÉVALUATION ET LA
PRÉVISION

2 Présentation du master SEP

Le Master 2 Statistique pour l'Évaluation et la Prévision (SEP) forme des statisticiens compétents et polyvalents, capables d'intervenir dans des domaines très variés. Il vise principalement deux objectifs : un objectif de savoir-faire qui consiste à procurer aux étudiants des compétences techniques et un objectif de polyvalence qui va leur permettre d'intégrer et de travailler au sein de toute structure dotée d'un service statistique et impliquant des interactions avec d'autres disciplines.

Une des spécificités de la formation est de confronter tout au long de l'année les étudiants à la pluridisciplinarité, ce qui permettra à ces derniers d'être capables de tenir un discours intelligible dans tous les domaines où les statistiques sont présentes

La formation des étudiants du master 2 se fait à travers des modules de cours bien définis dont on citera :

- L'analyse des systèmes complexes : L'analyse décisionnelle des systèmes complexes est une discipline qui vise à fournir des méthodes et outils de pilotage des systèmes complexes ou de modélisation des environnements chaotiques (théorie du chaos). La démarche insiste sur la transdisciplinarité et inscrit donc l'Analyse décisionnelle des systèmes complexes dans le courant de pensée systémique
- L'apprentissage automatique : L'apprentissage automatique ou apprentissage statistique est un champ d'étude de l'intelligence artificielle qui se base sur des approches statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, cela concerne la conception, l'analyse, le développement et l'implémentation de telles méthodes
- Les outils Big Data : Les bases de données relationnelles classiques ne permettent pas de gérer les volumes de données du big data. De nouveaux modèles de représentation permettent de garantir les performances sur les volumétries en jeu. Ces technologies, dites de Business Analytics & Optimization (BAO) permettent de gérer des bases massivement parallèles. Des patrons d'architecture (« Big Data Architecture Framework », BDAF) sont proposés par les acteurs de ce marché comme MapReduce créé par Google et utilisé dans le framework Hadoop. Avec ce système, les requêtes sont séparées et distribuées à des nœuds parallélisés, puis exécutées en parallèles (map). Les résultats sont ensuite rassemblés et récupérés (reduce). La famille de méthodes enseignées dans ce module comprend également le langage SQL (sigle de Structured Query Language, en

français langage de requête structurée) qui est un langage informatique normalisé servant à exploiter des bases de données relationnelles.

- L'Analyse des données et data mining : famille de méthodes statistiques dont les principales caractéristiques sont d'être multidimensionnelles et descriptives. Dans l'acception française, la terminologie « analyse des données » désigne donc un sous-ensemble de ce qui est appelé plus généralement la statistique multivariée. Certaines méthodes, pour la plupart géométriques, aident à faire ressortir les relations pouvant exister entre les différentes données et à en tirer une information statistique qui permet de décrire de façon plus succincte les principales informations contenues dans ces données.
- La gestion des risques, séries temporelles et économétrie.
- Les tests statistiques avancés sous R : Au contraire de la statistique descriptive, on va utiliser des
- lois de probabilités afin de prendre une décision dans une situation faisant intervenir une part de hasard
- Conférences, Business Intelligence, SAS et VBA.

A cette liste s'ajoute également des cours d'expression orale et de rédaction en Anglais, des séances d'aide à la prise de parole et les travaux d'Implication dans la vie universitaire (IVU).

3 Administration publique

La communauté d'agglomération du saint-Quentinois communément appelée Agglomération du Saint-Quentinois est une structure intercommunale française composée de 39 communes situées dans le département de l'Aisne en Région Hauts-de-France.

Lieu du stage : St Quintinois

Promotion 2018 :

Intitulé du stage :

Missions(s) : Alimentation de Bdd, référent de cette base avec mise en place d'un observatoire , chargé des études stratégiques de la direction. Veille stratégique, lien avec de nombreux services. Analyser les données note de synthèse, argumenter pour objectif d'aide à la décision.

Techniques statistiques :

Logiciels : R

3.1 Conseil Regional Nouvelle Aquitaine

Le conseil régional de Nouvelle-Aquitaine est l'assemblée délibérante de la région française de Nouvelle Aquitaine. Le conseil régional est composé de 183 conseillers régionaux élus pour 6 ans et est présidé par le socialiste Alain Rousset.

Lieu de stage : 14 Rue François de Sourdis 33070 Bordeaux Cedex

Promotion 2020 :

Intitulé du stage : Appui aux travaux de suivi des réalisations et de prospective

Missions(s) : Contribuer à la construction des tableaux de bords, indicateurs et à la mise en œuvre de la démarche de pilotage stratégique, au traitement et à l'analyse des données.

Techniques statistiques : Analyse des données, constructions de modèle prospectifs

Logiciels : R, Python, BO

3.2 Union départementale des centres communaux d'action sociale de la Seine-Saint-Denis (UDCCAS)

L'Union Départementale des Centres Communaux d'Action Sociale du département de la Seine Saint Denis (UDCCAS 93) voit le jour en 2003 à Saint Ouen par Bertrand DRUON (Maire Adjoint de Saint Ouen à l'époque). Les Centres Communaux d'Action Sociale (CCAS) naissent des cendres des Bureaux d'aide sociale par la loi n°86-17 du 6 Janvier 1986 à la suite des premières lois de la décentralisation. Les missions de ces organisations ont été fixées par un décret en mai 1995 (modifié en janvier 2000). A cet effet, elles ont sous leur administration la gestion des demandes de revenu de solidarité active, l'instruction des demandes de couverture médicale universelle et la participation aux actions sociales et Medico sociales de coordination

Lieu de stage : 93380 Pierrefitte-Sur-Seine

Promotion 2021 :

Intitulé du stage : Analyses des Besoins Sociaux Collectives en Seine Saint Denis

Mission(s) : Collecter des données sur le site de l'INSEE / Nettoyer les données / Calculer des indicateurs / Représentation graphique et cartographique des résultats Interprétation des résultats / Effectuer des entretiens pour la phase qualitative / Rédaction d'un rapport

Techniques statistiques : Enquêtes / Géomarketing

Logiciels utilisés : Rstudio / QGIS

4 Aéronotique, Espace, défense et sécurité

ArianeGroup, anciennement *Airbus Safran Launchers* (ASL), est une coentreprise créée en 2015 et détenue à parts égales par Airbus et Safran pour notamment développer les lanceurs Ariane 6. La société est chargée du développement et de l'intégration des lanceurs. Elle a plusieurs filiales dont Arianespace qui est chargée de la commercialisation.

Lieu de stage : Rue du Général Niox – BP 30056

Promotion 2020 :

Intitulé du stage : Ingénieur DATA Analyst DATA Scientist

Missions(s) :

- Découverte des outils, identification des données et de leur disponibilité
- Etude du besoin préalablement défini (indicateurs à créer via les extracts et croisement des données.
- Réalisation d'une maquette répondant au besoin est mis en oeuvre (conception de l'outil)
- Test et validation de l'outil

Techniques statistiques :

Logiciels :

4.1 Groupe Safran

Safran est un groupe international de haute technologie, leader dans les domaines de l'aéronautique, de l'espace, de la défense et de la sécurité. Il conçoit, développe, produit et commercialise des moteurs pour des avions civiles, des avions militaires et des satellites. En 2015, l'entreprise employait 70 100 personnes et a réalisé un chiffre d'affaire de 17,4 Milliards d'euros.

Lieu de stage : Rond point René Ravaud, 77550 Moisy-Cramayel

Promotion 2017 :

Intitulé du stage : Analyse des caractéristiques des aubes atypiques du moteur LEAP

Missions(s) : Compréhension des besoins métier des ingénieurs process /Analyse des données disponibles : collecter, explorer les données et vérifier leur cohérence /Identification de profils de pièces atypiques /identifier les causes racines potentielles expliquant ces profils atypiques, via des modèles d'apprentissage supervisé.

Techniques statistiques : K-Means, ACP, Interpolation de Lagrange, Clustering, ACM, AFDM, CAH

Logiciels : R

5 Banque et Assurance

Allianz IARD est une entreprise régie par le code des assurances créée en 1912, dont le siège social est situé à Paris. Elle est spécialisée dans le secteur des assurances. Elle emploie 2 000 à 4 999 salariés. Allianz France est un acteur majeur du secteur des assurances avec près de 13 milliards d’euros de chiffres d’affaires en 2018. Elle est classée dans la catégorie des grandes entreprises (GE).

Lieu de stage : 100-101 Terrasse Boieldieu –Tour Franklin, 92800 La défense 9

Promotion 2020 :

Intitulé du stage :

Missions(s) : Travail sur les contrôles internes des souscriptions des contrats d’assurance dans plusieurs branches différentes.

Techniques statistiques :

Logiciels : Pyspark, python, Jira, Git, Pycharm

5.1

AVIVA

Le groupe Aviva est un assureur généraliste qui propose tous types de produits d’assurances. C’est un des premiers assureurs Européens avec Axa, Allianz et Generali. De façon générale, ses missions se déclinent sur 3 principales activités qui sont : l’Assurance, la Protection et l’Epargne. Elles sont elles-mêmes divisées en sous-catégories : pour l’Assurance, on retrouve 4 catégories telles que l’Assurance Automobile, l’Assurance Habitation, l’Assurance Emprunteur et les Assurances Professionnelles. Pour la Protection, on retrouve 3 catégories telles que la Protection de la famille, la Prévoyance et la Santé (les complémentaires Santé). Pour l’Epargne, on retrouve également 3 catégories telles que l’Assurance vie, la Retraite et la Gestion financière. En clair, il faut noter que la partie Assurance consiste à assurer par le biais d’un contrat un ou plusieurs bien(s). Quant à la partie Protection, elle consiste à couvrir le salaire d’une personne indépendante lorsqu’elle a un accident de travail. L’Epargne, permet de réaliser des placements financiers.

Lieu de stage : 72000 Le Mans

Intitulé du stage :

Mission(s):

- Création, mise à jour, maintenance de tableaux de bords et d'indicateurs clés de performance
- Automatisation de tableaux de bord
- Construire des modèles de prévision
- Optimisation des traitements – Proposition d'axe d'amélioration
- Communication des résultats (reporting).

Techniques statistiques : Dashboard / Prédiction

Logiciels utilisés : Python / Excel / Google collab / Qlik

5.2 AXA France

AXA est un groupe français spécialisé dans l'assurance des personnes et des biens (véhicules, habitation, équipement), ainsi que dans la gestion des actifs. Le groupe couvre aussi bien les particuliers et leurs proches que les entreprises et leurs salariés.

Ce groupe est issu des diverses fusions et acquisitions de compagnies d'assurance dont la plus ancienne, l'Ancienne Mutuelle de Rouen, date de 1817. Le groupe devient le deuxième assureur français et prend le nom d'AXA en 1985. L'acquisition de l'Union des Assurances de Paris (UAP) en 1996 consacre AXA comme premier assureur de l'Hexagone et deuxième assureur mondial. AXA poursuit dans les deux décennies qui suivent son développement à l'international et dans les pays émergents. Le groupe AXA a ainsi été élu première marque d'assurance mondiale en 2017, et ce pour la neuvième fois consécutive. En 2021, AXA est le deuxième assureur européen, le quatrième assureur mondial en termes de CA (chiffres d'affaires) et le troisième en termes d'actifs non bancaires détenus.

Lieu de stage : Marly-le-Roi

Promotion 2022 :

Intitulé du stage : Stage en tant que Chargé d'études statistiques

Missions(s) : Concevoir et fabriquer des présentations des résultats de l'année 2021 en interne à l'attention des équipes commerciales et de souscription, et ce pour chacune des cinq régions géographiques en lesquelles est subdivisée la France. Les statistiques porteront sur le périmètre généraliste pour les salariés, les collectivités locales et les associations.

Techniques statistiques : Modélisation statistique

Logiciels : Word/ Excel/ Outlook/ Powerpoint/ WPS Analytics de la société World Programming/ SAS

5.3 BNP Paribas Cardif

BNP Paribas Cardif est un groupement d'intérêt économique du groupe BNP Paribas, spécialisée dans l'assurance. Créée en juillet 1973, la société est basée à Nanterre. En 2021, BNP Paribas Cardif emploie 8000 personnes, pour un chiffre d'affaires de 32,6 milliards d'euros.

Lieu de stage : 8 rue du Port, 92728

Promotion 2019 :

Intitulé du stage : Elaboration d'un score de rachat en assurance vie

Missions(s) : Transformer les sélections aléatoires de demandes de rachats vérifiés en une sélection ciblée sur les rachats anormaux.

Techniques statistiques : Régression logistique, forêts aléatoires

Logiciels : Python

5.4 Generali

Generali est une compagnie d'assurance italienne fondée en 1831. Elle fait partie des leaders de ce domaine, Il s'agit en effet de la 3ème compagnie au monde derrière Allianz et Axa. Elle est très largement implantée en France et une position forte sur l'assurance vie.

Lieu de stage : 11 Av. François Mitterrand, 93210 Saint-Denis

Promotion 2022 :

Intitulé du stage :

Missions(s) : La production de programmes de traitement et de mise à plat de données reçues par des fournisseurs extérieurs en vue de leurs ingestions dans le datalake Hadoop.

Techniques statistiques : Visualisation des données

Logiciels : SQL/ VBA

Promotion 2021 :

Intitulé du stage : Ingénieur de données

Missions(s) : Contribuer à la mise en œuvre d'une solution de traitements des données en vue de la création de reporting opérationnel des directions de marchés et du Comex/ Cadrer

les besoins du périmètre du service Pilotage Opérationnel./ Analyser les étapes de traitement de la donnée (ingestion, nettoyage, préparation, transformation, présentation, optimisation) / Participer à la mise en place d'un POC avec les outils proposés par la DSI · / Contribuer à qualifier les outils en fonctions des besoins du service Pilotage Opérationnel.

Techniques statistiques : Analyse des données

Logiciels : Dataiku/ Plateforme Big Data Hadoop-Spark

5.5 GROUPAMA

Groupama est une société d'assurance française. Elle possède deux marques commerciales :

- Groupama, marque d'assurance généraliste distribuée par le réseau des Caisses régionales. Elle est implantée dans 10 pays, notamment en Europe (France, Italie, Hongrie, Roumanie, Grèce, Bulgarie et Slovaquie), en Turquie, en Tunisie et en Chine ;
- Gan, assureur des entrepreneurs porté par Gan Assurances et Gan Eurocourtage, deux réseaux généralistes, ainsi que Gan Patrimoine et Gan Prévoyance, deux réseaux spécialisés.

Promotion 2020 :

Stage 1 :

Lieu de stage : 2 rue Léon Patoux - 51686 Reims cedex 2

Intitulé du stage : Data scientist

Missions(s) :

- Fiabiliser et retravailler la base de données sous R et/ou Python
- Appliquer les méthodes de data science pour analyser le profil des clients résiliés
- Formaliser les résultats obtenus

Techniques statistiques : Analyse de données, gestion des produits et des risques de l'assurance non-vie, modélisations statistiques, techniques de datascience, programmation sous R et Python.

Logiciels : Microsoft Office, R, Python

Stage 2 :

Lieu de stage : 8-10 rue d'Astorg, 75008 Paris

Intitulé du stage : Actuariat/ gestion des sinistres

Missions(s) : exploration de données, entretien, selection et validation des modèles.

Techniques statistiques : Machine learning, exploration de donnée, gestions de sinistres

Logiciels : Dataiku, R, Python

5.6 MGP

La MGP est une mutuelle d'assurance créée par et pour les policiers il y a plus de 60 ans. Il s'agit d'une structure ayant l'ambition de faire vivre un système favorisant l'accès à des soins de qualité pour tous ses adhérents, dans leur vie familiale mais aussi dans vie professionnelle.

Lieu du stage : 23 Rue de Palestro, 75002 Paris

Promotion 2017 :

Intitulé du stage : Tableau de bord et prédictions statistiques

Missions(s) : Mise à jour des tableaux de bord et consolidation de données /Réalisation d'une fiche de proposition d'action /Réalisation de prévisionnel

Techniques statistiques : Séries Chronologiques

Logiciels : Excel-VBA

5.7 BANQUE POPULAIRE VAL DE FANCE

La Banque Populaire Val de France est affiliée au groupe Banque Populaire et Caisse d'Épargne (BPCE) qui est le deuxième groupe bancaire en France. BPCE pratique une multitude de métier de la banque et de l'assurance en étant au plus près des besoins des personnes et des territoires, en s'appuyant sur ses deux grands réseaux coopératifs : Banque Populaire et Caisse d'Épargne, ainsi que sur ses filiales. Les 14 Banques Populaires détiennent des positions de premier plan auprès des PME, des professionnels et de la clientèle patrimoniale et les 15 Caisses d'Épargne auprès des particuliers, des collectivités locales et du logement social.

Lieu de stage : 78 180 Montigny le Bretonneux

Promotion 2021 :

Intitulé du stage : Réalisation Pilotage Commercial

Mission(s) :

- Mise à jLieu de stage:our des tableaux de bord du Pilotage
- Répondre à des analyses ponctuelles

Techniques statistiques : Etudes statistiques / Analyse de données / Tableaux de bords

Logiciels utilisés : Python, Hive, Spark, SAS

5.8

BNP PARIS BAS

La Banque Nationale de Paris (BNP), première banque française, créée en 1966 par les pouvoirs publics est issue de la fusion du Comptoir National d'Escompte de Paris (CNEP) et de la Banque Nationale du Commerce Internationale (BNCI). La Banque de Paris et des PaysBas devenue Paribas en 1982 est un autre établissement financier, qui, à la suite d'une bataille boursière, fusionne en 2000 avec la BNP pour donner naissance au groupe BNP Paribas.

Lieu de stage : 75009 PARIS

Promotion 2021 :

Intitulé du stage : Data Analyst

Mission(s) : Mener certains use-cases de leur définition jusqu'au delivery et à la passation aux équipes de production / Fournir des explications simples et taillées à la mesure des aptitudes techniques des personnes métiers sur les méthodes d'analyses de données et de data visualisations / Faire de la veille technologique, de la R&D et la lecture d'articles sur internet pour trouver les technologies inexistantes à l'interne et qui permettrait de résoudre d'éventuelles problématiques / Spécifier l'agrégation et le niveau de détail en fonction de la population à laquelle la restitution est destinée.

Techniques statistiques : Automatisation des données / Reporting / Tableau de bord

Logiciels utilisés : Python / SAS / R / domino

5.9 BNP PARIBAS personal Finance

BNP Paribas Personal Finance est n°1 du financement aux particuliers en France et en Europe au travers de ses activités de crédit à la consommation. Filiale à 100 % du groupe BNP Paribas, BNP Paribas Personal Finance compte plus de 20 000 collaborateurs et opère dans une trentaine de pays.

Lieu de stage : Paris

Promotion 2022 :

Intitulé du stage : Utilisation des données de navigation pour le profiling dans le e-marketing

Missions(s) : Travail sur un projet de marketing digital pour traiter des parties liées à la data science (comme le clustering) et des parties développement comme le regroupement de variables

Techniques statistiques : Machine learning

Logiciels : Python

5.10 BNP PARIBAS Securites Services

Lieu de stage : 9 rue du Débarcadère Pantin 93500

Promotion 2018 :

Intitulé du stage : Creation d'une interface permettant de gerer la gouvernance et la qualité des données

Missions(s) : Mise en œuvre du progiciel AB Initio pour gérer la gouvernance et la qualité des données. : Process de A à Z de data management sur un système d'info dédié.

Techniques statistiques : Data visualisation, Régression linéaire

Logiciels : Tableau

5.11 Caisse d'épargne Bretagne pays de loire

Implantée sur deux régions et neuf départements, la Caisse d'Epargne Bretagne Pays de Loire est une banque mutualiste régionale de proximité, solide et performante. Banque de référence pour les particuliers et les professionnels, elle est présente sur son territoire avec 420 agences et à distance grâce à Mon banquier en ligne. Partenaire majeur de l'économie locale, 16 centres d'affaires (Entreprises, Collectivités, Associations, Immobilier Professionnel, Logement Social et Economie Mixte) sont répartis sur tout le territoire. Avec ses filiales spécialisées, Sodero Gestion et Batiroc Bretagne Pays de Loire, la Caisse d'Epargne Bretagne Pays de Loire bénéficie également d'une expertise reconnue dans le domaine du capital investissement et du crédit bail immobilier.

Lieu de stage : 15 avenue de la jeunes seorvault 44700

Promotion 2018 :

Intitulé du stage : Analyse et modélisation des données du personnel

Missions(s) : Création de modèles statistiques à partir des données sociales pour faire de l'analyse prédictive. Construction d'un portail de reporting RH pour l'ensemble de la CEBPL SQL et R

Techniques statistiques : analyse predictive

Logiciels : SQL, R

Crédit Agricole Consumer Finance

Crédit Agricole Consumer Finance est une filiale de Crédit Agricole S.A, elle est issue de la fusion entre Sofinco et Finaref le 1er avril 2010. Son activité principale est la distribution de crédit à la consommation (Crédit classique, crédit auto et crédit renouvelable) via ses marques principales Sofinco, Viaxel, et Creditlift courtage, par les canaux de distributions (vente directe, financement sur le lieu de vente, e-commerce, partenariat et courtage). Elle propose aussi pour les particuliers des assurances et des produits pour la prévoyance contre les accidents et les aléas économiques.

Lieu de stage : 59100 ROUBAIX

Intitulé du stage : Conseil data

Mission(s) : Datamining / Machine Learning / Data Analyse / Modélisation mathématique et statistique exploitée en Data Science

Techniques statiques : Analyse de données / Datamining / Modélisation

Logiciels utilisés : Python / Hive / Spark / SAS / Office

5.12

Crédit Agricole

Leadeur de la banque de proximité en France et classé 10ème banque mondiale, le CA est engagé dans l'investissement socialement responsable et l'émergence d'un monde plus éthique, Il s'engage à la modernisation du réseau commercial et à l'essor de la banque digitale et poursuit activement son développement dans les métiers de gestion d'actif et de la banque privé en France, en Europe et à l'international. Toutefois, deux principes accompagnent le groupe depuis toujours dans son développement : l'utilité au travers des grandes transformations de la société et l'universalité de ses métiers, de ses offres, de ses territoires et de ses clients.

Lieu de stage : 92120 Montrouge

Promotion 2021 :

Intitulé du stage : Statistiques, contexte réglementaire en risque de crédit

Mission(s): Réalisation des backtesting des modèles internes (PD, LGD) permettant de s'assurer de la performance, la robustesse et du pouvoir prédictif des modèles internes bâlois / Migration des programmes SAS actuellement utilisés vers le logiciel R / Production de benchmarks permettant d'expliquer les éventuels écarts entre les notations internes et celles des agences externes de notation / Reporting de suivi des modèles internes / Travaux de

recherche sur de nouveaux indicateurs et tests statistiques applicables aux LDP (Low Default Portfolios).

Techniques statistiques : Backtesting / Automatisation des données

Logiciels utilisés : SAS / RStudio

5.13 Credit Agricole du Nord-Est

En France, le groupe Crédit Agricole est le premier acteur financier dans l'économie française. Il est considéré comme le Leader dans le domaine de la banque et l'assurance et il est la 1^{ère} banque de proximité avec plus de 20 millions de clients particuliers. Le Crédit Agricole se compose de 39 caisses régionales, qui se définissent comme étant « des banques coopératives qui assurent les fonctions commerciales, bancaires, financières et logistiques du Crédit Agricole » Chaque caisse régionale est indépendante et toutes partagent les valeurs mutualistes du groupe.

Lieu de stage : 25 rue Libergier 51088 Reims

Promotion 2022 :

Intitulé du stage : Typologie de clientèle & Score d'appétence

Missions(s) : La mise à jour d'une typologie de clientèle (36-49 ans) ainsi que la mise à jour d'un score d'appétence

Techniques statistiques : Machine learning

Logiciels : SAS/ ZEUS(déployée par CATS, dédiée au Big Data)

Promotion 2021 :

Intitulé du stage : Participer aux travaux de l'équipe DataMining

Mission(s) : Créer et/ou améliorer des scores, des typologies en intégrant l'exploitation de nouvelles données / Réaliser des études clientèle / Analyser les résultats obtenus, valider la cohérence des résultats.

Techniques statistiques : Segmentation / Etude de la demande / Scoring

Logiciels utilisés : /

Lieu de stage : 51088 REIMS CEDEX

Promotion 2020 :

Intitulé du stage : Scores et typologies dans le domaine bancaire

Missions(s) : Analyse descriptive, production de score d'appétence, de typologie de client

Techniques statistiques : Scoring, segmentation

Logiciels : SAS, Python, R

Promotion 2019 :

Intitulé du stage : Data mining au Crédit Agricole du Nord-Est: scores d'appétence et géomarketing

Missions(s) : Réalisation de scores d'appétence en flux, en crédit et en assurance... La méthode en flux consiste à regarder les profils des clients avant la souscription d'un produit, sur une période définie de 1 ou 2ans pour cibler les clients qui ont les mêmes profils aujourd'hui. Elle s'oppose à la méthode dite en <stock> qui consiste à regarder les profils des clients à un instant <t> pour cibler également les potentiels clients.....Création d'un datamart des données externes qui permettra de compléter des données internes de la banque dans la création de score ou typologie.....** Le datamart est un ensemble de données ciblées, organisées, regroupées et agrégées pour répondre à un besoin spécifique à un métier ou un domaine donné. Il est donc destiné à être interrogé sur un panel de données restreint à son domaine fonctionnel, selon des paramètres qui auront été définis à l'avance lors de sa conception....

Techniques statistiques : Credit scoring, géomarketing

Logiciels : Pack office, SAS (SAS Enterprise Guide, SAS Enterprise Miner)

Promotion 2018 :

Intitulé du stage : Typologie de clientèle et score d'appetence pour les clients sociétaires du crédit agricole

Missions(s) :

- Production de scores d'appétence
- Production de typologies de clientèle
- Analyses descriptives des sujets d'études
- Rédaction des rapports des études
- Mise en production industrielle
- Recherche et développement dans le domaine Big Data

Techniques statistiques : Clustering

Logiciels : SAS, R, Python

5.14 LE CREDIT LYONNAIS (LCL)

Depuis son rapprochement avec le Groupe Crédit Agricole SA en 2003, le périmètre d'activités de LCL, réseau national de banque de détail, est axé sur le marché des particuliers, des professionnels, des entreprises et la Banque privée.

LCL est une banque de proximité qui compte 2 065 implantations et 20 900 collaborateurs au service de 6 000 000 de clients particuliers, 320 000 clients professionnels et 27 000 clients entreprises et institutionnels.

Lieu de stage : 2 avenue de Paris, 94811 Villejuif Cedex

Promotion 2020 :

Intitulé du stage : Statistiques / Gestion des Risques / Data Sciences

Missions(s) : Construire de nouveaux algorithmes en testant des techniques innovantes liées à la data science, identifier et aider au pilotage du risque de crédit.

Techniques statistiques : Machine learning

Logiciels : Python, R, SAS, pack office

5.15 Oney Bank

Oney est une filiale d'Auchan Holding, spécialisée dans le crédit à la consommation, la monétique, la gestion des moyens de paiement et la connaissance des clients. Il sert 7,4 millions de clients en Europe et en Asie, dont millions en France. Oney a été créé en 1983 et est aujourd'hui présent dans onze pays (France, Portugal, Espagne, Italie, Pologne, Hongrie, Russie, Roumanie, Ukraine, Malte, Chine). En France, Oney banque regroupe 850 collaborateurs sur ses deux sites : le siège social à Croix et le site à Tours.

Lieu de stage : 34 Avenue de Flandre, 59 179 Croix

Promotion 2019 :

Intitulé du stage : Bonification du parcours client

Missions(s) : Optimisation de la segmentation de préqualification client partenaire à des fins marketing/risque (Analyse du comportement clients, simulation des apports opérationnels : amélioration des KPI (Key Performance Indicator) sur le parcours de souscription, ex: la part de client avec un parcours bonifié.....Etude de la performance de l'octroi et simulation de l'impact de l'assouplissement des règles d'octroi sur le taux d'acceptation.

Techniques statistiques : Segmentation, Modélisation statistique

Logiciels : SAS, Excel

Promotion 2017 :

Intitulé du stage : Analyse de la typologie des clients Oney et détection des niches à risque via une analyse de comportement

Missions(s) :

- Création d'indicateurs pouvant expliquer le risque client./L'exploitation de quelques données non utilisées jusqu'à aujourd'hui dans les segmentations effectuées par les membres de l'équipe.
- Analyse de la typologie des clients à l'aide des algorithmes de machine learning.
- Détection des niches risque via une analyse descriptive des clusters construits.
- Analyse des facteurs expliquant l'impayé du client et création d'un score de risque.
- Construction d'une vision 360° pour chaque comportement détecté.
- Lister les usages possibles de la classification pour permettre aux pôles métiers d'exploiter les résultats obtenus et de les intégrer dans leurs futures stratégies d'octroi ou de confiance client.

Techniques statistiques : Typologie clients/CRISP-DM/Custering Mixte (K-means, CAH, RLM)

Logiciels : SAS

5.16 RCI BANK AND SERVICES

RCI Banque SA, dont l'identité commerciale est Mobilize Financial Services (ex RCI Bank and Services), est implantée en France et dans 36 pays (principalement en Europe et en Amérique du Sud). Elle s'est développée autour du crédit automobile afin de permettre aux particuliers et aux entreprises d'acheter des véhicules du groupe. Elle distribue aussi d'autres services aux entreprises et aux particuliers, comme l'extension de garantie ou de l'assurance.

Lieu de stage : 92099 Paris La Défense Cedex

Promotion 2021 :

Intitulé du stage : Assistant Analyste Score

Mission(s) : Développement des modèles statistiques : Modèles de risque de crédit / Modèles de Data Mining Client, CRM, Marketing / Modèles de provisionnement / Garantie de la performance des modèles statistiques en production / Revue régulière de stabilité, performance et prévision de tous les modèles utilisés (Backtesting) et redéveloppement des modèles en cas de backtesting non satisfaisant / Développement d'un score d'octroi de crédit

Techniques statistiques : Analyse de données / Scoring / Modélisation

Logiciels utilisés : SAS / Excel / Python

Lieu de stage : 93160 NOISY-LE-GRAND

Promotion 2019 :

Intitulé du stage : Backtesting de la Probabilité de Défaut

Missions(s) :

- Analyser des modèles PD (la probabilité de défaut de l'emprunteur) et du processus backtesting existants;
- Veiller au respect des exigences réglementaires;
- Développer de nouveaux modules et tests dans le cadre du nouveau backtesting;
- Réaliser une documentation et une présentation du nouveau backtesting aux membres de l'équipe.

Techniques statistiques : Régression logistique

Logiciels : SAS, Excel

5.17 RCI banque -DIAC

RCI Banque est une société anonyme à conseil d'administration créée en 1980. En 1990, RCI avec DIAC (Diffusion Industrielle de l'Automobile par le Crédit). Aujourd'hui la DIAC est une filiale détenue à 100 % par RCI Banque et située à Paris. Elle se spécialise dans le secteur bancaire et emploie plus de 3700 salariés. En 2019, la banque a réalisé 2,1 milliards d'euros de chiffre d'affaires. Elle est classée dans la catégorie des grandes entreprises (GE).

Lieu de stage : 15 rue d'Uzès Paris

Promotion 2018 :

Intitulé du stage : Developpement de score marketing, comparaison de methodes de Scoring classiques vs Machine learning

Missions(s) :

- Comparaison des méthodes de Scoring classiques vs. Machine Learning avec utilisation potentielle de données externes.
- Mise à jour occasionnelle des monitoring & backtesting
- Analyses statistiques ponctuelles.

Techniques statistiques : Méthode de scoring, Machine learning

Logiciels : SAS, R

5.18 Société Générale

La société générale est un groupe bancaire français dont le siège social se trouve à la Défense à Paris. Elle a été créée en 1864 pour encourager le développement du commerce et de l'industrie en France. Elle est présente dans 67 pays à travers le monde. En 2019, le groupe possédait une réserve en capitaux propres de 63,5 milliards d'euros. La société générale appartient à la catégorie de grande entreprise avec plus de 10 000 salariés.

Lieu de stage : Paris

Promotion 2019 : SAS Spring Campus

Intitulé du stage : Segmentation analytique des clients CDN pour l'amélioration des alertes LCB-FT

Missions(s) : Mise en place d'une segmentation analytique(le clustering) des clients du Groupe Crédit Du Nord pour la détection des transactions pouvant relever du Blanchiment d'Argent et le Financement du Terrorisme

Techniques statistiques : segmentation client, machine learning, clustering

Logiciels : SAS Guide, SAS Visual Analytics, SAS AML 7.1, Hive,Pyhton, Talend, Apache Oozie, Jira.

Promotion 2018 :

Intitulé du stage : Integration du machine learning au sein des solutions de lutte contre la fraude et le blanchissement d'argent

Missions(s) :

- Implication dans les alertes générées par les scénarios.
- Définir, réceptionner les données nécessaires à la création des modèles et en vérifier la qualité.
- Définir et mettre en œuvre les méthodes d'échantillonnage.
- Créer les modèles de Machine Learning et documenter les travaux.
- Soumettre les modèles aux instances de validation et répondre à leurs recommandations.
- Analyser la performance des modèles.

Techniques statistiques : Machine learning, Apprentissage supervisé, regression, Random Forest

Logiciels : SAS

6 Commerce et Industrie

Adisseo est l'un des leaders mondiaux de la **nutrition animale** avec un chiffre d'affaires d'1.69 milliard d'euros en 2021 et plus de 3 900 clients. En 2006, l'entrée d'Adisseo dans le groupe Bluestar, lui a facilité l'accès au marché chinois, premier marché mondial en termes de croissance.

Lieu de stage : 10 Place du Général de Gaulle, 92160

Promotion 2020 :

Intitulé du stage :

Missions(s) :

- Aide à la collecte de données,
- Création de base de données,
- Traitement et analyse des données,
- Création plateforme de consultations des données des élevages

Techniques statistiques : analyse de données

Logiciels : Python, R, SQL

6.1 Arcelor Mittal

ArcelorMittal est un groupe sidérurgique mondial. Son siège social est installé à Luxembourg. En 2021, il est le deuxième plus important producteur d'acier au monde, avec 79,26 millions de tonnes produites. Il est classé 156 dans le classement 2016 Fortune Global 500 des plus grandes sociétés du monde.

Promotion 2021 :

Lieu du stage : 57280 Maizières-lès-Metz

Intitulé du stage : Supporter l'évolution des outils d'intelligence artificielle sur l'activité commerciale d'AMDS

Missions(s) :

- Analyser les données de nos datawarehouses actuel pour les transformer en axe d'info exploitable
 - Elaborer les critères de segmentation
 - Travailler en collaboration avec notre partenaire actuel afin :
 - Multiplier & formaliser les back test dans le but de valider les évaluations du Model
 - Monitorer et reporter les déviations du model, les analyser, afin de comprendre l'origine de cette déviation
 - Participer au workshop d'évolution des modèles en apportant une analyse technique
 - Finaliser la refonte de la proposition « minimum Price »
 - élaborer une prédiction du réel prix mini marché pour chaque offre de prix Europe
- Techniques statistiques : Analyse de données / Dashboard

Techniques statistiques :

Logiciels : Rstudio

Promotion 2019 :

Lieu du stage : 57280 Maizières-lès-Metz

Intitulé du stage : Modélisation statistique du LME(Liquid Metal Embrittlement)

Missions(s) : La mission principale est de tester de nouveaux types d'aciers pour le secteur automobile tout en répondant aux différentes préoccupations du secteur

Techniques statistiques : Régression linéaire multiple, Le modèle de régression logistique

Logiciels : Rstudio

6.2 BLT Cosmetics LABOTÉ

Blt Cosmetics Laboté est une entreprise évoluant dans le commerce de détail de parfumerie et de produits de beauté. Son effectif est compris entre 6 et 9 salariés. Sur l'année 2018 son chiffre d'affaires réalisé est de 204 700 euros. Elle appartient à la catégorie de petite et moyenne entreprise.

Lieu du stage : 11 rue madame 75006 Paris

Promotion 2020 :

Intitulé du stage : Analyse des données clients

Mission(s) : Proposer des methodes d'analyser les données clients, analyser et interpreter les données, proposer une algorithmes de prescriptions sur mesure

Techniques statistiques : Analyse de données, Gestion de projet

Logiciels : Excel

6.3 Chambre de commerce et d'industrie des Ardennes

La chambre de commerce et d'industrie d'ardennes est un organisme chargé de représenter les intérêts des 7139 entreprises commerciales, industrielles et de service du département des Ardennes et de leur apporter certains services. C'est un établissement public qui gère en outre des équipements au profit de ces entreprises.

Lieu de stage : 18A AVENUE GEORGES CORNEAU 08106 CHARLEVILLE-MEZIERES

Promotion 2017 :

Intitulé du stage : Etat, évolution et comparaison des établissements commerciaux ardennais de plus de 300m²

Missions(s) :

Techniques statistiques : Analyse des données, Clustering

Logiciels : Excel, R

6.4 PUM Plastiques

PUM Plastiques est le spécialiste français du négoce de produits et solutions plastiques à destination des Entreprises du Bâtiment et des Travaux Publics.

Lieu de stage : 4 rue René Francart, 51100 Reims

Promotion 2019 :

Intitulé du stage : Pricing, travaux d'analyses et d'études pour permettre une meilleure exploitation

Missions(s) : Mettre en place une refonte tarifaire, mise en place d'une base de données concurrence en relevant le prix tarif des articles chez les concurrents afin de suivre les évolutions, la détermination d'une remise unique par regroupement d'indice de cotation pour chaque segment (type de client, taille, région, type d'activité) pour ne pas provoquer de pertes.

Techniques statistiques :

Logiciels : Rstudio & Power BI , Excel

Promotion 2018 :

Stage 1 :

Intitulé du stage : Travaux de réflexions et de conception de l'évolution du tarif vente (entreprise internationale) : pricing

Missions(s) :

Techniques statistiques :

Logiciels : R Microstrategy, Excel

Stage 2 :

Intitulé du stage : Elaboration de reporting de suivi d'activités

Missions(s) :

- Accompagnement du service commercial dans le pilotage des activités à travers le suivi des performances :
- Etude et analyse du fichier clients grands comptes de la société;
- Pilotage de l'activité des données de clients avec la mise à jour de la base de données, la gestion et la structuration;
- Création et mise en place d'un processus d'automatisation des reportings de suivi d'activité;
- Mise en place d'un tableau de bord pour le suivi des performances commerciales (définition de KPIs et visualisation des données);
- Profilage des clients digitaux en fonction des différents canaux de vente.

Techniques statistiques : SGBDR

Logiciels : SQL, R, Excel

6.5 Zenpark

Zenpark est une société anonyme évoluant dans le secteur d'activité du commerce de gros (interentreprises) de fournitures et équipements divers pour le commerce et les services. Son siège se trouve à PARIS, elle appartient à la catégorie de petite et moyenne entreprise.

Lieu de stage : 142 rue Montmartre Paris

Promotion 2018 :

Intitulé du stage :

Missions(s) :

- Création d'applications sur Qlik : consulter les équipes métiers concernées, construire le modèle de données, importer les données, configurer les feuilles de visualisation des données
- Mise à jour des applications Qlik : suivre les changements en base de données SQL et les répercuter dans les applications Qlik concernées, prendre en compte les changements de process et les prendre en compte dans les applications Qlik concernées
- Smart-Cities : effectuer l'analyse de l'occupation réelle des parkings, créer un modèle de foisonnement applicable à différents types de parking
- Etude de rentabilité : créer un outil d'aide à la décision sur la performance des parkings
- Projection de revenus : créer un outil d'aide à la décision permettant de calculer des projections de revenus selon différents critères

Techniques statistiques : Data visualisation, modelisation mining

Logiciels : Qlik, SQL

6.6

7 Conseil

Créée en 2018, Algoan est une plate-forme SaaS qui propose une chaîne complète d'octroi de crédit. Algoan s'appuie sur les données de l'Open Banking pour prendre de meilleures décisions de crédit et garantir une expérience utilisateur fluide et transparente aux clients finaux. Algoan est une filiale à 100% de Yelloan, la Fintech française dédiée à l'accès au crédit pour les Millennials.

Lieu de stage : 20 Rue Drouot, 75009 Paris

Promotion 2022 :

Intitulé du stage : Machine learning appliqué à la détection des rejets de paiement dans le cadre de l'open banking

Missions(s) : Detection des rejets de paiements en créant une solution complète, innovante et automatisée

Techniques statistiques : Machine learning

Logiciels : Python

7.1 BGFİ Consulting

BGFİ Consulting est une société de technologies et service d'information qui a vu le jour en 2002. Elle est composée d'une cinquantaine de consultants spécialisées dans le domaine du Big Data et de l'Analyse. Elle compte plus que 200 clients pour lesquels, elle a pu élaborer plus de 400 projets. Au fil du temps, elle a une filiale au Canada basé à Vancouver (BGFİ NORTH AMERICA) et une autre en Tunisie (BGFİ Engineering).

Lieu du stage : 251 BD PEREIRE 75017 PARIS

Promotion 2017 :

Intitulé du stage :

Missions(s) :

- Extraction des qualifications correspondantes à des commentaires. (Création de fonctions de nettoyage de texte et de découpage des textes en mots et l'utilisation du web scraping)

- Détection des orientations des sentiments et de catégories d'opinions. (Création de dictionnaire des mots positifs et négatifs et utilisation de l'outil TreeTager pour la lemmatisation des mots, utilisation de quelques librairies de machine learning et d'analyse sémantique)

Techniques statistiques : Analyse textuelle

Logiciels : R, JSON

7.2 BI Consulting

BI consulting est une société de conseil qui évolue à Paris depuis 19 ans. Son activité principale est le conseil pour les affaires et autres conseils de gestion. Son effectif est compris entre 100 et 199 salariés, elle appartient à la catégorie de petite et moyenne entreprise. Sur l'année 2019, BI consulting a réalisé un chiffre d'affaire de 42 728 700 euros.

Lieu de stage : 142 Rue Montmartre, 75002 Paris

Promotion 2017 :

Intitulé du stage : Développement d'une application de rapprochement de faits d'actualité et de marchés financiers

Missions(s) :

- Conception d'une application de rapprochement entre des faits d'actualité et les marchés financiers pour prévoir l'impact d'un phénomène viral ou d'une rumeur provenant du web sur le cours d'une action
- Accompagner une grande banque française dans son choix d'un outil de Data visualisation

Techniques statistiques : Big Data, Classification, Régression logistique polynomique, Text-Mining, Data visualisation

Logiciels : Apache Spark, R

7.3 Business & Decision

Business & Decision est un groupe international de consulting et d'intégration de systèmes qui a forgé son expertise dans le domaine de la Business Intelligence (BI) et de la gestion de la relation client. Il reste un des acteurs de l'e-Business. Business & Decision offre des solutions adaptées à de nombreux secteurs d'activité : banque et assurance (Crédit agricole, La Banque Postale, BNP Paribas...), pharmacie, santé, Distribution et PCG (Unilever, DARTTY, LVMH,

LACOSTE...) , industrie (Sony, Audi, Alstrom..) , télécoms et médias (Canal+, M6, Bouygues Télécom, LeMondr,fr...), services publics et privés (LaPoste, APCE,...) etc...

Lieu de stage : Paris

Promotion 2020 :

Intitulé du stage : Etude économique sur des taux d'intérêt étrangers

Missions(s) :

- Faire un état de l'art sur le sujet du stage
- Identifier les différentes sources de données et la documentation nécessaire
- Collecter, trier et nettoyer les données identifiées comme pertinentes
- Faire une première analyse descriptive
- Créer différents jeux de données nécessaires à la modélisation
- Le cas échéant refaire un choix des variables pertinentes
- Mettre en œuvre différentes modélisations statistiques
- Comparer les modèles
- Conclure en répondant à la problématique posée (choix du modèle)

Techniques statistiques : Macro économie et finance, Modélisation statistique

Logiciels : SAS, R, Python

Promotion 2019 :

Stage 1 :

Intitulé du stage : Comparaison de données textuelles: étude de la similarité syntaxique entre les articles de presse

Missions(s) : Extraire des articles de presse en lignes, les analyser en vue de détecter les similarités (au sens syntaxique) entre elles.

Techniques statistiques : Modélisation (le web scraping : consiste à extraire des informations sur des pages web, et les mettre sous une forme exploitable. L'extraction s'effectue le plus souvent par des programmes informatiques par exemple les frameworks Python permettent de faire l'extraction.)

Logiciels : SAS & Python

Stage 2 :

Intitulé du stage : Etude des systèmes de recommandation: Cas d'usage et mise en œuvre pour des produits électriques

Missions(s) : Tester différentes méthodologies de recommandation, afin d'aider l'entreprise à mieux maîtriser les algorithmes de recommandations.

Techniques statistiques : KNN, scoring

Logiciels : SAS & Python

Stage 3 :

Intitulé du stage : Etude des déterminants du prix de l'immobilier en France

Missions(s) : Exploration de la base de données de la demande des valeurs foncières, publiée en Avril 2019 dans le but d'identifier les déterminants du prix de l'immobilier en France en se basant sur des méthodes statistiques et les approches de Machine learning

Techniques statistiques : Machine learning, Modélisation statistique

Logiciels : Python, Javascript

Stage 4 :

Intitulé du stage : L'intelligence artificielle pour la lutte contre la fraude des flammes publicitaires

Missions(s) : Utilisation des techniques de Deep Learning pour la Conception et entraînement d'un réseau de neurones convolutif pour la classification des images des plis en deux catégories: les plis avec flamme publicitaire et les plis sans flamme publication afin de détecter les machines. Cette classification va permettre de détecter les machines à affranchir qui utilisent les flammes publicitaires sans payer d'abonnement.

Techniques statistiques : Deep learning, computer vision, réseaux de neurones convulsifs, classification des images

Logiciels : Python, TensorFlow

Stage 5 :

Intitulé du stage : Génération automatique des observations à partir des réseaux sociaux : cas twitter

Missions(s) : La maîtrise des outils du text mining et la mise en place d'un algorithme d'extraction automatique de données sur twitter afin de créer un thésaurus propre à l'entreprise à partir d'une sémantique pré-définie..

Techniques statistiques : Extraction, Text mining, Thesaurus, Python, IA, observation, web scraping, machine learning

Logiciels : Python, Knime

Stage 6 :

Intitulé du stage : Modélisation du prix de l'immobilier

Missions(s) : Utilisation des techniques de Machine learning pour la prédiction des prix de l'immobilier en France en fonction des caractéristiques du logement. Mobilisation des connaissances statistiques et de modélisation linéaire multiple...Proof of Concept

Techniques statistiques : Machine learning, Data engineer, modèle explicatif, modèle prédictif (réseaux de neurones, arbres de décision, RandomForest), Data visualisation

Logiciels : Python, QGIS

7.4 Capgemini

Capgemini est le leader du marché de l'industrie du conseil et des services informatiques en entreprise dans le monde, sa croissance internationale s'est construite grâce à une stratégie de développement et de diversification de métiers. La société a été fondée par Serge Kampf le 1er octobre 1967 à Grenoble, en France, sous le nom de « Sogeti » (Société pour la gestion de l'entreprise et traitement de l'information). A l'instar de ses grands concurrents, le groupe Capgemini s'est développé à travers de plusieurs acquisitions dans tous les secteurs d'activités liés aux services informatiques : intégration de systèmes, l'infogérance, consulting, outsourcing.

Lieu du stage : 11 rue de Tilsitt 75017 Paris France

Promotion 2019 :

Intitulé du stage : Machine Learning dans le domaine de la Supply Chain: Cas du Groupe Carrefour

Missions(s) :

- Réécriture des codes pour l'Optimisation de leur temps d'exécution en vue de faire des prévisions des sorties d'entrepôt... [travail en projet via l'utilisation de données extraites d'un Data lake (Le Data Lake est un référentiel de données permettant de stocker une très large quantité de données brutes dans le format natif pour une durée indéterminée. Il regroupe les données structurées en provenance de bases de données relationnelles en couloir ou en colonne, les données semi-structurées telles que les CSV, les logs, les XML, les JSON, et les données non structurées telles que les emails, les documents et les PDF....Toutes les données d'une entreprise y sont stockées).
- Mise en place de modèle et prévision en utilisant les techniques de machine learning.

Techniques statistiques : Machine learning, forecasting, data lake, prévision

Logiciels : SAS : En particulier SAS Viya ; nouvelle plateforme de SAS hébergée dans le cloud. (Sas Studio, SAS Visual Analytics pour le reporting, SAS Visual Forecasting pour les prévisions et SAS Visual Data Mining and Machine Learning pour la préparation des données, l'extraction de leurs caractéristiques, la modélisation et la comparaison de modèles), Hive

Promotion 2017 :

Intitulé du stage : Global AML Transaction Surveillance

Missions(s) :

- L'intégration/restitution des données;
- L'analyse de la qualité des données;
- La rédaction de spécifications techniques détaillées (en Anglais);
- L'exécution de tests des développements réalisés,/A la coopération avec les équipes infrastructure et les autres services d' ITEC;
- La préparation à la mise en production.

Techniques statistiques : Datawarehouse

Logiciels : SAS

7.5 CERFRANCE

Cerfrance est une entreprise qui est répartie sur le territoire français (métropole et outre-mer) avec plus de 700 agence, et qui exerce depuis plus de 60 ans. Elle compte parmi elle près de 12 000 collaborateurs et plus de 320 000 clients. Elle a pour objectif d'apporter conseil et expertise comptable à ses clients. Pour Ce faire, ils garantissent à leur clientèle de la performance, des experts au service, une offre qui correspond au mieux à la demande, une possibilité de grande communication à travers un réseau développé.

Lieu du stage : 18 Rue de l'Armorique, 75015 Paris

Promotion 2022 :

Intitulé du stage :

Missions(s) :

- Développer le conseil en gestion de patrimoine en détectant les potentiels adhérents de Cerfrance Brocéliande qui auraient de l'appétence aux prestations de conseil en gestion du patrimoine.
- Mesurer la consommation des services par les adhérents via un tableau de bord.

Techniques statistiques : Visualisation des données / Machine learning

Logiciels : Python / powerbi

Promotion 2017 :

Intitulé du stage : Impact météorologique sur la clientèle

Missions(s) : Prédire au client, selon la météo, comment il va devoir appréhender sa saison financièrement, par l'estimation de l'impact selon sa sensibilité au climat.

Techniques statistiques : K-Means, CAH, ACP, Data visualisation

Logiciels : R, SPARK, Python, HTML, Sql

7.6 CGI

CGI, fondée en 1976 par Serge Godin est une des entreprises canadiennes de Services du Numérique (ESN) les plus importantes au monde. Elle est présente aujourd'hui dans 40 pays. En 2021, elle compte environ 80 000 professionnels, plus de 5500 clients et réalise un chiffre d'affaires de 8 milliards d'euros. En France, CGI figure parmi les 10 meilleures ESN avec un chiffre d'affaires de plus d'un milliard d'euros et plus de 11 000 professionnels

Lieu du stage : CGI France SAS – Immeuble Energies 3 avenue Belle Fontaine 35510 Cresson

Promotion 2022 :

Intitulé du stage :

Missions(s) :

- Missions interne à CGI: Etablir des modèles de deep learning via les réseaux de neurones pour détecter les objets sur les images et créer un Framework application permettant d'exposer les résultats de ces modèles.
- Mission externe en tant qu'administrateur SAS: faire en sorte que chaque utilisateur puisse accéder à l'application SAS et que ses droits d'accès aux données demandées soient bien définis.

Techniques statistiques : Deep learning

Logiciels : Python/ Flask/ Html/ Javascript/ CSS/ SAS

Promotion 2020 :

Intitulé du stage :

- Récouter des données pour la création d'un réseau de neurones

- Création d'une application sous FLASK utilisant le réseau et la méthode d'industrialisation choisie

Techniques statistiques : Deep learning

Logiciels : SAS Viya, SAS Base 9.4M5, Windows, Linux, Python

7.7 CROSS SYSTEMS

Cross est une ESN qui dépend totalement de la Holding Micropole Suisse, qui elle-même dépend totalement du groupe Micropole France. L'ESN propose des prestations en conseils et de services en systèmes d'information.

Lieu de stage : 1227 GENEVE LES ACACIAS

Promotion 2021 :

Intitulé du stage : Mise en place d'une solution data de bout en bout

Mission(s) : Mettre en place les plateformes Cloud Data Platform (Snowflake, GCP, AWS, Azure) / Mise en place d'un ou plusieurs ETL pour l'intégration des données / Mise en place d'un ou plusieurs outils de restitution (Qlik Sense, Tableau Software ...) / Choix et mise en place d'une solution de machine.

Techniques statistiques : Automatisation des données / Benchmarking / Modélisation / Tableaux de bord

Logiciels utilisés : Snowflake, GCP, Azure, AWS, différents outils ETL, Qlik Sense, Tableau software, Power BI

7.8 Data Company

Data company est une entreprise de consulting créée en 2019 dont le siège social se trouve à paris. Son activité principale est le conseil en systèmes et logiciels informatiques. Son effectif est compris entre 3 et 5 salariés. Sur l'année 2019, son chiffre d'affaires s'élève à 4404700 euros. Elle appartient à la catégorie petite et moyenne entreprise.

Lieu de stage : 32 RUE de paradis 75010 Paris

Promotion 2020 :

Intitulé du stage : Analyse des données clients

Mission(s) :

- Explorer les Open data liées à la consommation des français pour trouver des axes de comparaison avec les données transactionnelles fournies par nos 60 partenaires
- Analyser et recommander des secteurs de consommation forts ou émergents (Par exemple les services à la personne)
- Identifier des faiblesses à corriger dans la structure de la composition de la mégabase
- Massifier les historiques de performance de campagnes marketing pour déterminer des segments par type d'offres et de canaux exploités avec une ouverture vers une dimension globalisée sur le territoire français
- Construire les bases d'une prédiction de réactivité de sollicitation à des dons aux organisations caritatives, selon les données géomarketing.

Techniques statistiques : Machine learning, scoring

Logiciels utilisés : R

7.9 Data Vectis

Data Vectis est une société de service spécialisée dans le secteur d'activité du conseil en systèmes et logiciels informatiques. Elle est en activité depuis 2018 et offre des formations en big data et ses technologies en collaboration avec un cabinet spécialisé. Les activités de Data Vectis couvrent également les prestations suivantes : Conseil et assistance à la maîtrise d'ouvrage les transformations digitales et l'optimisation des processus métiers de même que le support aux grands projets de transformation

Lieu du stage : 19 Rue de Montyon, 75009 Paris

Promotion 2022 :

Intitulé du stage : Data Engineer au PMU / Data scientsit en interne

Missions(s) :

- Maintenir le data lake et le faire évoluer. Migrer les projets existants de OnPremises vers le data lake AWS.
- Automatiser la récupération des informations présentes dans un CV

Techniques statistiques : Visualisation des données

Logiciels : Python(Text Mining et de Natural Langage Processing (NLP))/ Scala, Spark /AWS Cloud/ Dataiku/ Git/ Confluence

Promotion 2019 :

Stage 1 :

Intitulé du stage : Réalisation d'une solution Big data/Machine learning de monitoring d'un système d'information

Missions(s) : Réalisation d'une solution Big Data et machine learning pour la détection, la classification et la prédiction des erreurs dans un système d'information des imprimantes connectées.

Techniques statistiques : Naive bayes (est un algorithme de classification multiple qui suppose que tous les attributs des données d'apprentissage sont indépendants. Il calcule la proba conditionnelle de chaque attribut.), régression logistique (qui est utilisée pour développer un modèle de régression lorsque la variable cible ou prédictif est catégorique. Cet algorithme exige que les attributs et la variable cible soient linéairement indépendants.), et enfin l'arbre de décision (il associe une proba à chacun des choix possibles en fonction du contexte de la décision).

Logiciels : Spark, Kibana (visualisation), HDFS (Hadoop Distributed File System : est un système de fichiers distribué conçu pour fonctionner sur du matériel standard, tolérant aux pannes et convient aux applications disposant de grands ensembles de données et assure un accès à haut débit aux données), Kafka (est une plateforme Apache distribuée de streaming qui construit un pipeline entre des systèmes et des applications qui doivent avoir accès au même flux de données.)

Stage 2 :

Intitulé du stage : Estimation de la redevance de circulation des trains

Missions(s) : Mise en place de différents cas d'utilisation du Produit Minimum Viable (MVP) au travers d'un pipeline de transformation des données au sein de l'écosystème Big data.....Le MVP est la stratégie développement d'un produit répondant au besoin métier par toutes ses fonctionnalités.. Il gère l'estimation à partir des commandes de la redevance de circulation des trains

Techniques statistiques : big data, regression

Logiciels : Spark

7.10 DSI-Consulting

DSI consulting est une société à responsabilité limitée localisée à LE PLESSIS-ROBINSON (92350) depuis 7 ans. La société est spécialisée dans le secteur d'activité du conseil en systèmes et logiciels informatiques. Elle emploie entre 10 et 19 salariés, et appartient à la catégorie de petite et moyenne entreprise. Son chiffre d'affaires s'élève à hauteur de 1 267 600 euros sur l'année 2019.

Lieu du stage : 41 Avenue du Général Leclerc, 92350 Le Plessis-Robinson

Promotion 2017 :

Intitulé du stage : Conception et développement d'une solution environnement agricole avec semences climat

Missions(s) :

Techniques statistiques : Arbre de décision, KNN

Logiciels : Spark, Python, R

7.11 Efficience 3

Fondé à Reims en 1985, Efficience3 est un institut d'études marketing et opinion indépendant. Fort d'une croissance approfondie des marchés et clients, actif dans plus de 80 pays et territoires, l'entreprise relève beaucoup de défis et possède une équipe qui s'implique avec simplicité, convivialité et professionnalisme dans chacun de ses projets. Depuis 30 ans, l'institut d'études indépendant d'une cinquantaine de salariés implanté à Reims réalise des sondages, enquêtes d'opinion, études qualitatives, en France et en Europe, pour le compte de clients internationaux, en garantissant service client et respect de la qualité (ISO 9001). Lorsqu'Efficience3 fut créée, elle était sous forme d'association, la société telle qu'elle existe aujourd'hui est constituée en janvier 1985 avec une activité essentiellement consacrée aux études et au conseil à caractère commercial auprès des entreprises dans la région du Nord-est. Ce n'est qu'à partir de 1992, que la SARL se positionne sur les études marketing et entame un développement international. En 2002, l'entreprise se diversifie par le lancement d'un Omnibus qui lui permet d'aborder de nouveaux marchés.

Lieu du stage : 26 Rue Buirette, 51100 Reims

Promotion 2017 :

Intitulé du stage : Mise en place des études téléphoniques au sein du département CATI & Online

Missions(s) :

- Apporter assistance et aide aux assistants et chargés d'études internes dans la mise en place des études téléphoniques et online. Ceci passe par les différents paramétrages, les tests, les recherches d'informations préalables...
- Exploiter les résultats, en passant par la vérification de données, rédaction et relecture des rapports.
- Mettre à jour et optimiser une macro VBA
- Mettre en place un outil pour la segmentation : réflexion sur les méthodes et son application à différentes études.

Techniques statistiques : ACM, CAH

Logiciels : R, Excel VBA, Askia

7.12 Equancy

Equancy est un cabinet de conseil international, basé à Paris et Dubai, spécialisé dans la transformation digitale des entreprises. Son activité principale consiste à la planification, la conception et la mise en œuvre des solutions Big Data, Data Science et Intelligence Artificielle pour ses clients.

Lieu du stage : 4 Rue Jules Lefebvre, 75009 Paris

Promotion 2022 :

Intitulé du stage : Identifier de nouvelles techniques de monitoring en machine learning et étudier la mise en œuvre de ces techniques sur un modèle de Churn développé et déployé en production en 2019 pour La Poste Mobile.

Missions(s) : L'optimisation de la stratégie de ciblage de Nespresso et l'identification des nouvelles techniques de monitoring de modèles de Machine Learning en production.

Techniques statistiques : Machine learning

Logiciels : SPSS Modeler, SQL, BigQuery, GCP, Oracle, Python, PySpark, DataBricks, Azure

7.13 Greenflex

Créée en 2009, GreenFlex est devenue l'un des principaux acteurs européens de son secteur avec plus de 600 clients. GreenFlex accompagne ses clients dans leur gestion efficace de l'énergie en combinant les métiers de l'intelligence des données et de la gestion et du financement d'équipements.

Lieu du stage : 16 boulevard Montmartre Paris

Promotion 2018 :

Intitulé du stage : Mise en œuvre opérationnelle du CRM en mode multicanal pour la coopération CRM Services

Missions(s) :

- Mise en œuvre opérationnelle du CRM en mode multicanal pour la coopération CRM Services (requête sur UNICA - logiciel maison- ciblage ; mise en place projets et tests ; mise en production ; enrichissement; bilan quanti et quali ; développer des outils

- Collecte, structuration et enrichissement de données (via des API)
- Datamining et exploration des bases de données : analyses statistiques (sous R)
- Définition et validation d'indicateurs
- Interprétation des résultats par croisement avec l'expertise métier
- Élaboration des outils de reporting (sous QLIK)

Techniques statistiques : Data mining, Data visualisation

Logiciels : R, Qlik Sense

7.14 Greenomy

Greenomy est une startup Regtech (Regulatory – Technology) fournissant aux entreprises, aux investisseurs et aux prêteurs des outils réalisables et rentables pour automatiser la génération, la validation, la gestion et la divulgation des données ESG qui sont nécessaires pour se conformer au règlement de l'UE sur la Taxonomie ainsi qu'aux réglementations et normes ESG actuelles et futures (par exemple, SFDR, NFRD, CSRD, etc).

Lieu de stage : Luxembourg

Promotion 2021 :

Intitulé du stage : ESG & Sustainability Data Analysis

Mission(s) :

1. Mise en œuvre de la taxonomie et soutien à la conformité :
 - Synthétiser et traduire les informations complexes de la taxonomie de l'UE en documents clairs, informatifs et convaincants ;
 - Mise en correspondance entre les réglementations dans le cadre du plan d'action de l'UE sur le financement de la croissance durable (SFDR, NFRD) et au niveau international (TCFD, SDGs, SASB, GRI, etc) ;
 - Recherche des politiques, des tendances et des meilleures pratiques liées au calcul des critères et des seuils pertinents pour les activités éligibles ;
 - L'évaluation préliminaire des activités économiques (rigueur et utilité).
2. Analyse des données :
 - Travailler en collaboration avec l'équipe informatique pour mettre en œuvre les concepts de taxonomie .

- Participer à l'implémentation et l'analyse des données
3. Suivi et politiques :
- Participer à la durabilité de la roadmap de Greenomy.
 - Taches ad-hoc (par exemple, préparation de présentations, réunions, documentations)
 - Évaluer les opportunités du marché et faire des recommandations.

Logiciels utilisés : Bloomberg, CDP, RStudio, Slack

7.15 Groupe AVISIA

Le Groupe AVISIA est un cabinet de conseil, intégration et réalisation de projets data. Il est organisé en 3 pôles que sont Conseil et AMOA, Décisionnel et Big Data et Performance Analytique. En tant que SAS Platinum Partner, l'entreprise propose une offre complète de services autour de la Data en restant ouverte et experte sur toutes les technologies modernes utilisées en Data Science.

Promotion 2018 :

Lieu du stage : 48 Avenue Victor HUGO Paris

Intitulé du stage : Mise en place d'un projet d'alimentation de la base extranet d'AVISIA contenant l'ensemble des descriptifs des missions et des compétences des consultants d'AVISIA

Missions(s) :

- Compréhension des besoins et des enjeux du projet.
- Réflexion sur la réalisation du projet.
- Automatisation de l'extraction des informations contenues dans les CV des collaborateurs au format Word.
- Réalisation de phases de test et de validation.
- Création de bases de données associées.
- Réalisation de tests pour l'alimentation des bases extranet.
- Validation du programme.

Techniques statistiques : SAS, SQL, R, VBA

Logiciels : Python

Promotion 2017 :

Lieu du stage : 8 Avenue Kléber, 75116 Paris

Intitulé du stage : Etude Digital Analytics et Reporting : cas du e-commerce en B2B

Missions(s) :

- Concevoir et réaliser une architecture des outils associés à la stratégie digitale ;
- Mettre en place des processus d'automatisation de la collecte, de la transformation et de la restitution des données ;
- Concevoir et réaliser un tableau de bord des données online et offline

Techniques statistiques : Analyse d'audience

Logiciels : Google Data Studio, Python, R

7.16 Hanalytics

Hanalytics accompagne les entreprises dans le développement de leur activité grâce à la data. Ses employés mettent leurs compétences en Data & Digital au service des entreprises afin qu'elles puissent activer et exploiter dans les meilleures conditions leur capital Data. Hanalytics intervient à tous les niveaux de la chaîne de valeur de la Data.

Lieu du stage : Paris

Promotion 2022 :

Stage 1 :

Intitulé du stage : Data gouvernance : Cas de Valorissimo

Missions(s) :

- Mission 1 : Valorissimo Analyses comportementales des utilisateurs de la plateforme ; Product Analysis ; Market Analysis ;
- Mission 2: Création de pipeline de données et des pipelines d'analyse de données avec SQL et Python

Techniques statistiques : Visualisation des données / Analyse des données

Logiciels : Google Cloud Platform/ Data Build Tools/ Google Data Studio/ Google Analytics/ Google Tag Manager/ Deepnote/ Secoda/ Airbyte/ Stitch/ Hightouch/ Metaplane-Data

Stage 2 :

Intitulé du stage : Analyse des facteurs qui influence le taux de clic sur les bannières promotionnelles

Missions(s) : Analyse des facteurs qui influencent le taux de clic sur les bannières promotionnelles pour chacun des clubs (club Marmara, club Lookéa et Nouvelle frontière) de la compagnie Tui France

Techniques statistiques : Machine learning

Logiciels : Google analytics/ DBT/ BigQuery/ Rstudio

Stage 3 :

Intitulé du stage : Modern Data Stack – Data Quality

Missions(s) : Faire des reportings des Kpi's de l'entreprises de façon hebdomadaire, généralement les lundi au profit de l'équipe Marketing de Mes demoiselles de Paris. D'autres demandes s'ajoutent pour le même client telles que la mise en place des pipelines de données sur SQL à des fins d'analyse, la mise en œuvre de certains algorithmes de machine Learning sur python (Scoring, segmentation...), la refonte des Dashboards et leur maintenance sur Data Studio et la documentation reposant sur la description des variables, et des tables des différentes bases de données grâce à l'outil Secoda.

Techniques statistiques : Visualisation des données / Machine learning

Logiciels : Python/sql

7.17 Innogence Consulting

« Innogence » comme Innovation et Intelligence, c'est un cabinet de consulting spécialisé en Intelligence économique et en stratégie d'innovation. L'intelligence économique c'est l'ensemble des activités coordonnées de collecte, de traitement et de diffusion de l'information utile aux acteurs économiques, en vue de son exploitation. Basé en France et en Côte d'Ivoire ; Elle accompagne ses partenaires et clients à mieux maîtriser leur environnement.

Lieu de stage : 93100 Montreuil

Promotion 2021 :

Intitulé du stage : Chargée d'étude de marché

Mission(s): Conception d'une démarche méthodologique et choix des outils les plus pertinents pour des études de marché / Collecte de données quantitatives et qualitatives dans divers secteurs d'activités, notamment sur le marché africain / Analyse et synthèse des résultats des études de marché / Élaboration de rapports d'études de marché (français et anglais) / Présentation des résultats obtenus et formulation de recommandations stratégiques à l'issue des études / Veille concurrentielle et technologique

Techniques statiques : Analyse de données

Logiciels utilisés : Excel / R studio / Xmind

7.18 JANASENCE

Janasense est un service connecté de prévention des risques de fragilité de la personne pour une sérénité partagée du couple aidant/aidé. Des révélateurs de vie équipés de capteurs permettent, avec une dose d'intelligence artificielle, d'identifier et d'anticiper les défaillances et risques de fragilité par une analyse prédictive et donc d'accompagner le mieux-être de la personne. Et se servir de toutes ces données pour des buts médicales après : des symptômes liés à la dépression ou encore une mauvaise qualité nutritionnelle.

Lieu du stage : 1 Avenue du Champ de Mars, Le Lab'O, 45100 Orléans

Promotion 2017 :

Intitulé du stage : Modélisation des comportement

Missions(s) : Construire des modèles de machine learning afin de modéliser les comportements et détecter les de la vie quotidienne de la personne telles que : se réveiller, dormir, prendre un bain, etc. Aussi, je suis amenée à mettre en œuvre des modèles afin d'identifier les zones d'anomalies et les variations de rythme de vie nominal au sein du logement.

Techniques statistiques : Modèle bayésien, modèle linéaire, Série temporelle, CAH, ACP

Logiciels : R,SPSS

7.19 Kheopsys Lab

Lieu du stage : Paris

Promotion 2022 :

Stage 1 :

Intitulé du stage : Mise en place d'un MLOps sur Cloud

Missions(s) : Concevoir un MLOps en développant un pipeline Machine

Learning sur la plateforme Vertex AI au niveau de Google Cloud Platform.

Techniques statistiques : Machine learning MLOPS

Logiciels : Google Cloud Platform et Amazon Web Service /Python/ Slack c/Asana

Stage 2 :

Intitulé du stage : Industrialisation des pipelines Machine Learning

Missions(s) : Mettre en place un mécanisme qui sert à automatiser les tâches d'importation, de traitement des données, ainsi que le training du modèle Machine Learning et la prédiction. Deux pipelines sont à implémenter : Un pipeline de training qui sert à entraîner le modèle de prédiction et l'évaluer. L'output final sera un fichier Excel contenant les identifiants et les prédictions correspondantes. On pourra également visualiser la performance du modèle en affichant l'accuracy, le score AUC, la courbe ROC, le F1 score ainsi que la matrice de confusion. Le deuxième pipeline est un processus industrialisé et adapté à la production. Il traite les nouveaux flux de données, et génère des prédictions en utilisant le modèle déjà entraîné et stocké par le premier pipeline. Le processus s'achève par la configuration d'un mécanisme CI/CD1 pour assurer l'intégration continue des modifications apportées sur les deux pipelines.

Techniques statistiques : MLOPS

Logiciels : google Cloud Platform/ python/ BigQuery

7.20 QUANT AI LAB

QUANT AI Lab est une firme internationale, spécialisée dans la Data, l'IA et l'analytique avancée pour l'audace, la performance et la valeur.

Promotion 2022 :

Intitulé du stage : Détection des anomalies

Missions(s) : Procéder à l'implémentation de techniques pour la détection des outliers au sein de QBox.

Techniques statistiques : Machine learning

Logiciels : Stramlit/python

7.21 Umanis SA

Umanis, créée en 1990 par Laurent Piepszownik, était une entreprise de services numériques française active dans l'informatique décisionnelle et dans les métiers liés aux mégadonnées.

Lieu du stage : 5 rue Georges de Loup Lyon

Promotion 2018 :

Intitulé du stage : Automatisation des transferts de fichiers : Transformation de fichiers XML en tables SAS

Missions(s) : Intervenir sur la maintenance applicative des applications SAS d'un des clients grand compte : 80% pges SAS sous UNIX, 20% pg shell ; tests unitaires et d'intégration ; analyse des demandes d'évolutions ; livraison en recette.

Techniques statistiques :

Logiciels : Windows/Linux, SAS Enterprise Guide, Putty, WinSCP

7.22 Velvet Consulting

Velvet Consulting est le cabinet de conseil leader dans l'accompagnement des organisations dans leur orientation client. Depuis 2004, grâce ses 250 passionnés de Marketing, Velvet accélère la performance des entreprises en construisant une expérience client des plus complètes et efficaces. L'approche globale et la complémentarité des expertises du cabinet Velvet, accompagnent la transformation centrée client, data et digitale des entreprises. Des orientations stratégiques à la mise en œuvre opérationnelle et technologique.

Lieu du stage : 64 rue la Boétie Paris

Promotion 2021 :

Intitulé du stage : Data Science, Analyse des parcours digitaux et modèles d'attribution

Missions(s) : Echange avec les différents acteurs médias présents au sein du groupe WPP et en prenant en main l'ensemble des données relatives à cet environnement (données de navigation, données publicitaires, planning media, etc.). Le projet se concentrera ensuite sur la mise en place des modèles d'attribution et évoluera selon l'avancement du stagiaire, les problématiques rencontrées et résultats obtenus. Selon l'avancement du stage et le profil du candidat retenu, une phase de développement d'interface web pourrait être à prévoir afin de mettre à disposition ces modèles et permettre l'analyse du rôle des différents média dans la conversion

Techniques statistiques : Dashboard / Analyse de données / Modélisation / Création d'application

Logiciels : R / Python

Promotion 2018 :

Stage 1 :

Intitulé du stage : Réalisation d'une étude de faisabilité dans le domaine de gestion des revenus (pour le compte d'une entreprise de service)

Missions(s) : Optimisation des actions marketing :

- analyses sur le parcours client ;
- segmentations marketing ;
- modélisation prédictive ;
- automatiser et maintenir les tableaux de bord de pilotage d'activité ;
- études

Techniques statistiques : machine learning

Logiciels : Python

Stage 2 :

Intitulé du stage :

Missions(s) : Mise en place d'un système de recommandations :

- Contribution à la conception et la mise en place des briques applicatives
- Développement de pipelines algorithmiques devlpt et implé. Des algo de recommandations.
- Développement en environnement big data (Spark, Hive, Python,) sur des pbtiques d'intégratonde données et d'industrialisation du machine learning.

Techniques statistiques : machine learning, series temporelles

Logiciels : Spark, Hive, Python

8 Eau et Assainissement

Entreprise centenaire et actuellement leader mondial des services d'eau et d'assainissement, de la gestion des déchets et de l'optimisation des services énergétiques, Veolia conçoit et déploie des solutions intégrées et adaptées aux nouveaux enjeux de ses clients et aux défis des ressources. Le groupe Veolia, référence mondiale en termes de gestion optimisée des ressources, est présent sur les cinq continents avec plus de 163 000 salariés. Au travers de ses trois activités complémentaires que sont la gestion de l'eau, de l'énergie et des déchets, Veolia contribue à développer l'accès aux ressources, à préserver les ressources disponibles et à les renouveler.

Lieu du stage : 291 Avenue Dreyfous Ducas

Promotion 2017 :

Intitulé du stage : Stratégie de localisation des clients pour le ramassage des déchets industriels

Missions(s) :

- Étude bibliographique concernant la collecte des déchets industriels
- Elaboration de l'architecture d'analyse des données massive du fait de la grosse capacité de données
- Analyse descriptive des données afin d'extraire le maximum d'informations pertinentes /Visualisation graphique des coordonnées GPS
- Recherche des bonnes stratégies au moyen d'algorithmes pour identifier la localisation des clients à partir des logs GPS des camions de collecte

Techniques statistiques : Partitionnement des données, K-médoïdes

Logiciels : R

9 Economie et développement

L'Organisation de coopération et de développement économiques (OCDE) est une organisation internationale qui œuvre pour la mise en place de politiques meilleures pour une vie meilleure fondée en 1961 à Paris. Elle est spécialisée dans le secteur d'activité de l'économie et le développement. Son effectif est compris entre 250 à 499 salariés et compte 35 pays.

Lieu du stage : 2 rue André Pascal, 7577 5Paris Cedex 16

Promotion 2017 :

Intitulé du stage :

Missions(s) : Harmoniser les données collectées des différents services économiques et le système d'information utilisé pour compiler et publier les résultats statistiques

Techniques statistiques : calcul d'indicateurs, statistique descriptive

Logiciels : SAS, R, SQL-Server, Excel

9.1 Centre d'études de politiques pour le développement (CEPOD)

Le Centre d'Etudes de Politiques pour le Développement (CEPOD) est un projet de renforcement des capacités placé sous la tutelle du ministère de l'Economie et des Finances du Sénégal. Il bénéficie de l'appui financier de la Fondation pour le Renforcement des Capacités en Afrique (ACBF) qui est située à HARARE, au Zimbabwe. Le CEPOD a été mis en place officiellement le 09 octobre 2003 par Monsieur Abdoulaye DIOP, Ministre de l'Economie et des Finances du Sénégal, en présence de Monsieur Soumana SAKHO, Secrétaire exécutif de l'ACBF et de son Directeur Monsieur Aliou FAYE.

Lieu du stage : Dakar

Promotion 2018 :

Intitulé du stage : Etude des facteurs déterminants des investissements directs étrangers dans les pays de l'Union économique monétaire ouest africain (UEMOA)

Missions(s) :

- Cadre théorique et Conceptuel de l'étude des IDE

- Analyse sectorielle et origine des flux d'IDE
- Approche théorique et empirique (Revue de la littérature sur les déterminants des IDE)
- Analyse de la relation économétrique entre l'IDE et les différentes variables qui le déterminent le modèle
- Données et résultats de régression
- Conclusions et recommandations de politique économique

Techniques statistiques : Régression linéaire

Logiciels : R

10 Edition de logiciels

SAS Institute est une multinationale américaine spécialisée dans les logiciels d'analyse, dont le siège est situé à Cary, en Caroline du Nord. SAS développe et commercialise un ensemble de logiciels d'analyse (également connus sous le nom de SAS) permettant d'accéder à des données, de les gérer, de les analyser et d'établir des rapports afin de faciliter la prise de décision. L'entreprise est la plus grande entreprise privée de logiciels au monde et ses logiciels sont utilisés par la plupart des entreprises du classement Fortune 500.

Lieu du stage : Domaine de Grégy- Grégy sur Yerres- 77257 BRIE COMTE ROBERT CEDEX

Promotion 2021 :

Stage 1 :

Intitulé du stage : Consulting Data Science – Fraude/Risk

Missions(s) : Qualification du risque client / Création de scénarios de détection/ Création de workflows permettant de matérialiser les processus utilisateurs / Création et personnalisation des Interfaces Utilisateurs (Ecrans et rapport de restitution) / Manipulation des données (Contrôle DQ, Profilage ...)

Techniques statistiques : Mapping / Classification / Machine learning

Logiciels : SAS

Stage 2 :

Intitulé du stage : Associate Business Operations

Missions(s) : Rédiger des rapports, identifier, analyser et coordonner les informations requises par les différentes unités business pour soutenir leurs plans d'amélioration continue et pour aider à déterminer les solutions aux problèmes business, développer, mettre à jour et maintenir régulièrement un référentiel central des informations de l'entreprise chez SAS, en l'utilisant comme une plate-forme, définissant les meilleures normes et partage des connaissances et fournissant des ressources pour faciliter et améliorer la valeur et la qualité de la livraison des projets.

Techniques statistiques : Prédiction / Tableaux de bords

Logiciels : SAS/ Excel

Promotion 2020 :

Stage 1 :

Intitulé du stage : Traitement d'image médicale et deep learning

Missions(s) :

- Traitement d'image médicale et deep learning.
- Maîtrise des algorithmes de Machine Learning et Deep Learning,
- Maîtrise de méthodes statistiques et de la modélisation prédictives

Techniques statistiques : Algorithmes de Machine Learning et Deep Learning, méthodes statistiques et modélisation prédictives

Logiciels : R , Python , SAS

Stage 2 :

Intitulé du stage :

Missions(s) : La mission adressera une ou plusieurs thématiques liées aux prévisions de sorties des entrepôts. Plus précisément, le rôle du stagiaire sera de prendre en charge un sujet précis (qui sera déterminé en fonction de l'avancement du projet au moment du stage), afin de le développer puis de l'intégrer dans le processus d'amélioration des prévisions

Techniques statistiques :

Logiciels : SAS VIYA – SAS Visual Analytics

Stage 3 :

Intitulé du stage : Consultant PreSales DATA MANAGEMENT

Missions(s) : Comprendre les problématiques business, proposer de nouvelles approches

Techniques statistiques : analyse predictive

Logiciels : SAS, Python, Javascript, Powerpoint

Promotion 2019 :

Stage 1 :

Intitulé du stage : Cycle de vie du modèle: Industrialisation du pilotage des indicateurs de validation et de performance du modèle

Missions(s) : La mission est de participer au projet de vente de la solution MRM (Model Risk Management). Cette solution est conçue afin d'aider les établissements financiers à gérer leur risque de modèle. Elle prend en charge le cycle de vie complet du modèle et crée un ensemble de rapports afin d'assurer le suivi et le contrôle de toutes les catégories de risques liés aux

modèles. Cela permet de répondre aux demandes du régulateur en fournissant des rapports réglementaires.

Techniques statistiques : Modélisation statistique

Logiciels : SAS

Stage 2 :

Intitulé du stage : Intégration du machine learning au sein de la solution SAS de lutte contre le blanchiment d'argent

Missions(s) : Ce stage combine connaissances d'entreprise et connaissances techniques. Les activités réalisées dans le cadre de ce stage vont de la préparation des rendez-vous et des séminaires et(formation dédiés aux clients SAS à la création des démonstrations logicielles. Elles font appel aux techniques de machine learning pour mettre en place et tenir à jour un algorithme de détection des cas de blanchiment d'argent au sein d'un système bancaire

Techniques statistiques : Blanchiment d'argent, segmentation, cluster, K-means, Classification hiérarchique ascendante, machine learning

Logiciels : SAS (SAS Enterprise Miner, SAS Enterprise Guide). Le premier simplifie le processus de data mining pour développer rapidement des modèles, comprendre les relations clés et identifier les schémas déterminants... Le deuxième est orienté vers la construction de requêtes (extraction de données, jointures) et la statistique descriptive mais possède aussi quelques interfaces pour la construction de modèles

Promotion 2018 :

Stage 1 :

Intitulé du stage : Intégration du machine learning au sein des solutions de lutte contre la fraude et le blanchiment d'argent.

Missions(s) :

- Comprendre la problématique business
- Rechercher et analyser les approches existantes
- Proposer de nouveaux modèles
- Développer & intégrer ces modèles
- Valider la valeur ajoutée de la nouvelle approche
- Présenter les résultats

Techniques statistiques :

Logiciels : R

Stage 2 :

Intitulé du stage :

Missions(s) :

(techniques de modélisation, open sources, big data...) et formaliser des recommandations pour un cycle de vie des modèles optimisé ; démonstrations logicielles, présentations clients, réponses à appel d'offres

Techniques statistiques :

Logiciels : R

Stage 3 :

Intitulé du stage :

Missions(s) :

- Mettre en pratique les nouvelles techniques de traitement de l'image et les illustrer à partir d'un démonstrateur sur cette problématique métier.
- Participer à des présentations internes ainsi qu'à des opportunités clients sur le même sujet.

Techniques statistiques : Machine Learning. Manipulation et traitement des données massives (Big Data)

Logiciels : SAS® Visual Data Mining & Machine Learning. SAS (SAS® Viya) et Jupyter.

Stage 4 :

Intitulé du stage :

Missions(s) : préparation de propositions pour valider l'adaptabilité des logiciels SAS pour les scénarios IoT et autres « edge computations » (traitement de flux de données en temps réel et Big Data, par exemple) en utilisant les techniques et technologies les plus récentes.

Techniques statistiques :

Logiciels : R

Stage 5 :

Intitulé du stage :

Missions(s) :

- Analytics Pre-Processing et démonstration de la qualité des outils SAS sur des données réelles basée sur la fraude aux soins médicaux à Bruxelles, Belgique.
- Réalisation d'un programme accessible permettant la création des variables RFM sous SAS Viya
- Participation à des présentations et des démonstrations des solutions SAS
- Travail et collaboration avec une équipe internationale

Techniques statistiques :

Logiciels : R

Stage 6 :

Intitulé du stage :

Missions(s) : Travailler au sein de l'équipe projet sur les différentes phases de mise en oeuvre d'une solution SAS de lutte contre la fraude, le blanchiment d'argent et l'abus de marché

Techniques statistiques :

Logiciels : R

Stage 7 :

Intitulé du stage :

Missions(s) :

- Participerez à l'ensemble des activités liées à l'avant-vente d'un projet (présentations, réponse à des appels d'offres, création de démonstrateurs, etc.).
- Création de contenus de Dataviz' dans un contexte où l'Open Data vient compléter les données disponibles chez nos clients innovants

Techniques statistiques :

Logiciels : SAS Visual Analytics, SAS Enterprise Guide, SAS Viya et Base, Forecast Studio, Visual Investigator

11 Energie

ATMO Grand Est a été créé en 2016 à la suite de la fusion des AASQA (Association Agrée de Surveillance de la Qualité de l’Air) d’Alsace, de Champagne-Ardenne et de Lorraine. Elle est une association agréée par le ministère de la transition écologique, qui surveille la qualité de l’air selon une approche transversale air-climat-énergie-santé, dont les principales missions répondent aux règles relatives à la Loi sur l’Air et l’Utilisation Rationnelle de l’Energie (Loi LAURE) du 30 décembre 1996. Cette approche résulte de l’interdépendance entre les quatre enjeux air, climat, énergie, et santé, dont l’un impact les autres, par exemple plus on consomme/ produit de l’énergie

Promotion 2022 :

Lieu de stage : 9 Rue Marie Marvingt, 51100 Reims

Intitulé du stage : Exploitation de données satellitaires

Missions(s) : Etudier les données MERRA2, données mises à disposition par la NASA concernant les particules fines. Le but était de les exploiter pour établir un lien entre les données satellitaires et les mesures au sol qui s’expriment dans des grandeurs physiques différentes.

Techniques statistiques : Machine learning

Logiciels : R studio / Python / (Qgis)

Promotion 2021 :

Lieu de stage : 67300 Schiltigheim

Intitulé du stage : Modélisation et prévision de la qualité de l’air dans la région Grand Est

Mission(s) : Comprendre le fonctionnement du modèle de dispersion et prendre en main les jeux de données associés / Réaliser une recherche bibliographique sur l’état de l’art des techniques de corrections statistiques / Développer une correction statistique performante des données brutes / Valider la méthodologie statistique en s’appuyant sur des scores (biais, corrélations...) / Tester la sensibilité du post-traitement et affiner sa paramétrisation.

Techniques statistiques : Prédiction / Modélisation / Analyse de données / Scoring Logiciels utilisés : RStudio / Python / QGIS

11.1 Akuo Energy SAS

Lieu du stage : 140 Avenue des Champs Elysées 75008 PARIS

Promotion 2020 :

Intitulé du stage : Wind Performance Analysis

Missions(s) :

- récupération des données de supervision des éoliennes via API de Bazefield
- analyse de l'intégrité de ces données
- création d'alertes et de rapports quotidiens sur la performance des parcs éoliens
- développement de modèles pour piloter la maintenance prédictive du groupe
- présentation des résultats à une équipe internationale (préventive et corrective)/gestion de projet international

Techniques statistiques : analyses statistiques sur des séries temporelles/algorithmes de machine learning (classification, regression, clustering, forecasting)/connaissance détaillée du fonctionnement d'une éolienne/compréhension des différentes stratégies de maintenance (préventive et corrective)/gestion de projet international

Logiciels : Windows 10/SQL Server/Python/Jupyter NoteBooks (ou autre)/Bazefield (logiciel de supervision)

11.2 EDF

En activité depuis plus de 65 l'Electricité de France (EDF) est localisée à Paris (siège social). Son activité principale est la production de l'électricité. EDF emploie plus de 10 000 salariés, elle se classe parmi la catégorie des grandes entreprises. Sur l'année 2019 son chiffre d'affaires réalisé est de 46 155 000 000 euros.

Lieu de stage : 33700 MERIGNAC

Promotion 2019 :

Intitulé du stage : Data science pour la maintenance du réseau de distribution HTA souterrain

Mission(s) : Etablir un modèle fiable et rigoureux qui permettra de déterminer les départs du réseau souterrain à risque de tomber en panne afin de prioriser les renouvellements préventifs. Plus précisément, estimer la probabilité de défaillance d'un départ l'année suivante en fonction des contraintes météorologiques et de l'historique de pannes subies.

Techniques statistiques : Cross-validation , SVM, arbre de décision , KNN , régression logistique et ACP.

Logiciels utilisés : Python

11.3 ENEDIS

Anciennement appelé ERDF (Electricité Réseau Distribution France), Enedis est une entreprise de service public. Elle est une filiale d'EDF (Electricité de France) et est spécialisée dans le secteur d'activité de la distribution d'électricité. Elle s'occupe de la gestion, de l'exploitation et la modernisation du réseau électrique en France métropolitaine et continentale. Enedis contribue activement à accompagner les territoires dans leur dynamique de transition énergétique et la production d'énergie renouvelable. Le dépannage 24h/24h, le relevé des compteurs, les raccordements des clients et toutes interventions techniques représentent le quotidien des salariés d'Enedis.

Lieu de stage : 33700 MERIGNAC

Promotion 2021 :

Intitulé du stage : développement d'application sous Power BI

Mission(s) : Le stagiaire devra s'approprier une ou plusieurs applications développées sous Qlickview pour les développer sous Poser BI, notre nouvel outil de data visualisation tout en optimisant la mise à disposition des données.

Techniques statistiques : Transformation de la donnée / Tableaux de bord

Logiciels utilisés : Power BI, Office

11.4 ENGIE SA

Engie est un groupe industriel énergétique français dont l'activité principale est le commerce de combustibles gazeux par conduites. L'Etat français est son actionnaire majeur, qui détient un quart de son capital. Le groupe est issu de la fusion de Gaz de France et Suez en 2008. Son siège social se trouve à la Défense et compte entre 2000 et 4999 salariés. Son chiffre d'affaires en 2019 s'élève à 17 282 000 000 euros. Elle appartient à la catégorie des grandes entreprises.

Lieu du stage : 1, place Samuel Champlain 92 La Défense

Promotion 2021 :

Intitulé du stage :

Missions(s) :

- Analyse de sentiment sur des données issues du web
- Mise en valeur par la visualisation des résultats d'analyse et automatisation des reportings
- Analyse de datasets pour fiabiliser le catalog de données en cours de réalisation
- Participation à des projets data en ML, NLP ou autres suivant les besoins

Techniques statistiques : Analyse de données structurées et non structurées

Logiciels : SQL, Python, Jupyter Notebooks, AWS S3 Platform, PowerBI

Promotion 2019 :

Lieu du stage : Paris / Paleseau

Intitulé du stage : Data science pour la maintenance du réseau de distribution HTA souterrain

Missions(s) : Etablir un modèle fiable et rigoureux qui permettra de déterminer les départs du réseau souterrain à risque de tomber en panne afin de prioriser les renouvellements préventifs. Plus précisément, estimer la probabilité de défaillance d'un départ l'année suivante en fonction des contraintes météorologiques et de l'historique de pannes subies.

Techniques statistiques : Cross-validation (il s'agit d'une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage : par exemple "tests et validation", "k-fold cross-validation" et "leave-one-out cross-validation"), SVM(le modèle SVM est un puissant discriminant. Il est souvent utilisé lorsque l'on veut affecter des individus à une classe donnée. Le principe de l'algorithme se base sur la recherche du meilleur hyperplan qui pourra séparer au mieux les données et maximiser la distance entre les classes.), arbre de décision (qui est une famille de modèles non paramétriques supervisés.), KNN (ce modèle permet de classer les individus selon les k plus proche voisins.), régression logistique et ACP.

Logiciels : Python

11.5 Kayros

Kayros est la principale société de géo-analyse avancée de l'énergie et de l'environnement qui aide les commerçants, les investisseurs, les opérateurs et les gouvernements à prendre de meilleures décisions. Kayros tire de la valeur de l'intégration de données alternatives et de marché dans des solutions uniques et des offres de produits spécifiques aux clients tout en mesurant l'impact environnemental et en fournissant un aperçu des risques liés au climat et à la transition énergétique.

Lieu du stage : Paris

Promotion 2019 :

Intitulé du stage : Détection de changement en imagerie radar

Missions(s) : La mission est de réaliser la détection de la phase de fracturation hydraulique dans la chaîne d'exploitation du pétrole et notamment du passage des camions, nécessaires à la réalisation de cette phase

Techniques statistiques : Séries temporelles

Logiciels : Python

11.6 PETROSEN TRADING & SERVICES

PETROSEN est la société des pétroles du Sénégal qui est majoritairement détenue par l'Etat sénégalais (99%). Créée en mai 1981, elle est sous l'autorité du ministère du Pétrole et des Énergies. La société s'occupe de la recherche et l'exploitation des hydrocarbures du sous-sol, mais aussi du raffinage jusqu'à la commercialisation. Elle a pour objectif de rendre le pays autonome en hydrocarbures et de sécuriser son approvisionnement.

Lieu du stage : 104 Sotrac Mermoz, Route de Ouakam, Dakar, Sénégal

Intitulé du stage : Budgeting, planning et veille stratégique dans l'aval pétrolier

Missions(s) : budgetting, valorisation et analyse des prospects pour informer la prise de décision , veille stratégique

Techniques statistiques : valorisation des actifs

Logiciels : MS365, ERP

11.7 VEOLIA

Entreprise centenaire et actuellement leader mondial des services d'eau et d'assainissement, de la gestion des déchets et de l'optimisation des services énergétiques, Veolia conçoit et déploie des solutions intégrées et adaptées aux nouveaux enjeux de ses clients et aux défis des ressources. Le groupe Veolia, référence mondiale en termes de gestion optimisée des ressources, est présent sur les cinq continents avec plus de 163 000 salariés. Au travers de ses trois activités complémentaires que sont la gestion de l'eau, de l'énergie et des déchets, Veolia contribue à développer l'accès aux ressources, à préserver les ressources disponibles et à les renouveler.

Lieu du stage : 291 Avenue Dreyfous Ducas

Promotion 2017 :

Intitulé du stage : Stratégie de localisation des clients pour le ramassage des déchets industriels

Missions(s) :

- Étude bibliographique concernant la collecte des déchets industriels
- Elaboration de l'architecture d'analyse des données massive du fait de la grosse capacité de données
- Analyse descriptive des données afin d'extraire le maximum d'informations pertinentes / Visualisation graphique des coordonnées GPS
- Recherche des bonnes stratégies au moyen d'algorithmes pour identifier la localisation des clients à partir des logs GPS des camions de collecte

Techniques statistiques : Partitionnement des données, K-médoïdes

Logiciels : R

11.8 Voltalis

Voltalis est une entreprise privée créée en 2006 spécialisée dans la flexibilité électrique. Sa mission principale consiste en l'ajustement en temps réel des flux d'électricité afin de répondre au besoin d'équilibre électrique selon la règle : production=consommation. En d'autres termes, c'est la capacité d'un site électrique à faire des économies d'énergie selon la demande.

Lieu de stage : 75008 Paris

Promotion 2021 :

Stage 1 :

Intitulé du stage : Pilotage budgétaire de la consommation électrique par prévision court et moyen terme

Missions(s) : le stagiaire et son encadrant détermineront au sein de cycles de 3 semaines des objectifs intermédiaires pour former un programme de travail concerté. Au sein de chacun de ces cycles, les résultats des analyses menées par le stagiaire (extraction BDD, traitement des données, analyses, modélisation, algo, mise en place de POC, etc.) devront permettre d'alimenter la réflexion et amener à déterminer les objectifs du cycle suivant. Le stagiaire est un membre à part entière de l'équipe et participe donc à ses rituels périodiques (principalement journalier, hebdo et 1 fois par cycle de 3 semaines).

Techniques statistiques : Modélisation / Prédiction

Logiciels utilisés : R Studio, Python

Stage 2 :

Intitulé du stage : Désagrégation du signal de consommation électrique des logements

Mission(s) : analyse des courbes de charge, estimation des postes de consommations électriques, analyse des variables explicatives de la consommation comme les variables météo ou les caractéristiques du logement, satisfaction adhérent, pragmatisme, compromis théorie/pratique

Techniques statistiques : Analyse de données / Modélisation / Tableaux de bords

Logiciels utilisés : RStudio / Python

Lieu de stage : 75008 Paris

12 Grande distribution

Auchan est une enseigne de grande distribution faisant partie de l'Association familiale Mulliez. L'enseigne est fondée en 1961 et emploie 354 851 salariés à travers le monde. Son chiffre d'affaires en 2018 est de 50,3 milliards d'euros. Elle est considérée comme une grande entreprise (GE).

Lieu du stage : 200 rue de la recherche – 59650 Villeneuve d'Ascq

Promotion 2020 :

Intitulé du stage : Amélioration de la connaissance clients AUCHAN

Missions(s) :

- Segmentation des clients en utilisant l'ensemble des datas disponibles (interne / externe)
- Optimisation du contenu de nos campagnes marketing personnalisées : systèmes de recommandation établis via la modélisation prédictive et le machine learning
- Définition des cibles les plus pertinentes pour les plans d'animation commerciales (scoring...)

Techniques statistiques :

Logiciels : Google Cloud Plateform, Teradata, R / Python, Microsoft Excel

12.1 Siège mondial Carrefour

Lieu du stage : Massy Palaiseau Paris

Intitulé du stage : Chargée des données cobotiques

Missions(s) :

- Inclure les données des prestataires dans le reporting data cobotique hebdomadaire sur python.
- Rédaction des synthèses et comptes-rendus d'événements importants.
- Création d'une documentation des projets en incluant les brochures et cartes visite associées.

Techniques statistiques : Visualisation des données

Logiciels : Python / stramlit

13 Hôtellerie

Evian Resort est une société active depuis 63 ans dont le siège se trouve à EVIAN-LES-BAINS. Son activité principale est l'hôtellerie et hébergement similaire. Elle appartient à la catégorie des grandes entreprises. Sur l'année 2019 elle réalise un chiffre d'affaires de 64 962 800 euros.

Lieu du stage : Evian

Promotion 2019 :

Intitulé du stage : Mise en place de l'outil CRM (Customer Relation Management) et Analyse de délai de réservation de la clientèle hôtel de l'evian resort

Missions(s) : Mise en place d'un DataLake à partir de données extraites depuis plusieurs SGBD (Un Système de Gestion de Base de Données) et analyse du comportement de réservation de la clientèle pour la mise en place d'un modèle de prédiction des réservations

Techniques statistiques : Extraction de données, prévision, marketing

Logiciels : R, Java

14 Informatique et Télécommunication

BIG APPS est une société de services informatiques. Elle intervient auprès des PME et de grands comptes pour la mise en place de solutions innovantes qui permettent de tirer pleinement profit des solutions et technologies autour de la Big Data, ainsi que de la data science afin d'optimiser le capital immatériel que représentent les données.

Lieu du stage : Romainville

Promotion 2018 :

Intitulé du stage :

Missions(s) : Etudier et implémenter un système d'interrogation d'un data lake basé sur le langage humain comme système de requête. La restitution de la donnée sera sous forme graphique (machine learning, process NLP).

Techniques statistiques :

Logiciels : R

14.1 Data moove

Data Moove est une start-up créée en 2014 qui propose une gamme complète de produits technologiques innovants pour le BtoBtoC (business to business to consumer) dans le domaine du tourisme : la plateforme Big Data sémantique City Moove collecte de l'offre touristique à sa diffusion (City Moove) et son exploitation avancée sur tout média digital tel que application mobile, site web, réseaux sociaux ou via l'assistant touristique digital Minotour.

Lieu du stage : 245 route des Lucioles 06560 Sophia Antipolis, France

Promotion 2019 :

Intitulé du stage : Intelligence artificielle et web sémantique pour l'assistance digitale

Missions(s) : Ledit stage fait appel à diverses techniques d'intelligence artificielle notamment dans le domaine du web sémantique. Dans le cadre de ce stage, le stagiaire est amené à contribuer à la mise en place d'un système d'assistance automatisé au service de la clientèle (les touristiques) à travers le développement de collecteurs et la construction de nouvelles bases de connaissances. Une base de connaissance, regroupe un ensemble connaissances spécifiques

à un domaine spécialisé donné , sous une forme exploitable par un ordinateur. Ces données (connaissances) proviennent de sources variées et sont récupérées grâce à des collecteurs.

Techniques statistiques : web sémantique, base de connaissance, machine learning, deep learning. Le web sémantique encore appelé le web de données permet aux machines de comprendre la sémantique, la signification de l'information sur le web. Il fournit un modèle qui permet aux données d'être partagées et réutilisées entre plusieurs applications, entreprises et groupes d'utilisateurs.

Logiciels : Python, RASA, JAVA. RASA est une plateforme qui permet aux développeurs de créer, d'améliorer et de déployer de meilleurs assistants (Intelligence artificielle) alimentés par une source ouverte.

14.2 Line Up Data Science

LINEUP DATA SCIENCE est une startup française créée en 2016 à Paris. Elle se spécialise dans le secteur d'activité du traitement de données, hébergement et activités connexes. Elle emploie 90 salariés. Elle a réalisé un chiffre d'affaires de 5685215 d'euros en 2018. Elle se situe dans la catégorie des petites et moyennes entreprises (PME).

Lieu du stage : 7 rue de Caumartin 75009 PARIS

Promotion 2020 :

Intitulé du stage : Data Science orienté Marketing

Missions(s) :

- Cadrage du besoin avec le client (cas d'usage associé, période observée, pilotes envisagés, apprentissages attendus : théorie, programmation, outils, méthodologie Data Science, etc.)
- Etat de l'art et mise en évidence de méthodes adaptées aux problématiques
- Analyse des outils, langages les plus pertinents pour mettre en œuvre les modélisations à construire

Techniques statistiques :

Logiciels : AWS, Dataiku, GCP, Python, SQL

14.3 Linkfluence FRANCE

Linkfluence est une startup française spécialisée dans l'écoute et l'analyse du web social créée en 2006. Elle est rachetée en 2021 par l'entreprise américaine Meltwater.

Lieu du stage : 5 Rue choron 75009 Paris

Promotion 2020 :

Intitulé du stage : Science des données appliquée en environnement "big data"

Missions(s) : analyser les données sociales, explorer les nouveaux algorithmes de machine learning, construire des visualisations, fournir une expertise aux équipes de vente, de marketing et de R&D

Techniques statistiques : Regression, NLP, clustering, séries temporelles

Logiciels : Radarly, Python, Pyspark

14.4 Monstock

Monstock est une société qui fournit des solutions de gestion de stocks créée en 2018 à Reims. Elle est spécialisée dans le secteur d'activité de la programmation informatique. Son effectif est entre 3 à 5 salariés. Son capital social est évalué à 13570 euros. Elle appartient à la catégorie petite et moyenne entreprise (PME).

Lieu du stage : 17 Rond-point de l'Europe, 51430 Bezannes

Promotion 2020 :

Intitulé du stage :

Missions(s) :

- Participer à construire des scores prédictifs (performances supply chain, stockage, picking,... etc)
- Segmentation clustering pour mise en place d'une approche type Magic Quadrant gartner
- Mettre en place une procédure de qualification de la data
- Dashboarding & data visualization via une plateforme type Looker
- Structuration des tables big query, documentation, conception de dictionnaire de donnée
- Mettre en place l'industrialisation des dashboarding avec une plateforme type Qlik sense
- Suivre le pilotage du pipeline data associé

Techniques statistiques : Gestion de projet, Data qualification, Data visualization

Logiciels : R, SQL, Python, AWS Aurora, DynamoDB, Redshift DB React, Python, Apache, Jupyter, Gitlab S3

14.5 Mybestpro (telecommunications)

Mybestpro est une entreprise active depuis 16 ans évoluant dans le secteur d'activité du conseil en systèmes et logiciels informatiques dont le siège social se trouve à PARIS. L'entreprise emploie entre 100 à 199 salariés, elle appartient à la catégorie de petite et moyenne entreprise. Son chiffre d'affaires sur l'année 2019 s'élève à hauteur de 16 399 500 euros.

Lieu du stage : Paris

Promotion 2018 :

Intitulé du stage :

Missions(s) : Algo pour système de recommandation, text mining avec prédiction de taux d'acceptation (sur document de type devis), analyse d'impact.

Techniques statistiques :

Logiciels : R

14.6 Orange SA

Orange Business Services (OBS) est une entité du groupe Orange, elle fournit des services informatiques et de télécommunications dans le monde, à destination des clients professionnels. C'est une entité commerciale que le groupe a constitué le 1er juin 2006. Elle est présente dans 220 pays et territoires et propose des services de connectivité aux entreprises. En quelques chiffres, OBS offre ses solutions informatiques à plus de 2 000 000 de clients professionnels PME dans le monde, plus de 100 000 client Cloud Pro et plus de 3 000 multinationales. Pour assurer le bon fonctionnement des solutions proposées ainsi que l'accompagnement client. L'entreprise a embauché plus de 21 000 collaborateurs dédiés aux activités entreprises. C'est grâce à cela que l'entreprise assure un chiffre d'affaires de 6,4 milliards d'euros en 2016.

Promotion 2022 :

Lieu du stage :

Intitulé du stage : Stage en tant que Data Analyst

Missions(s) : Projet R3l2 regroupant les ventes des reconditionnés, des reprises, des recyclés, du leasing et des locations. Chargée des locations et de l'uniformisation des différents suivis: -

Recréer le datamart des locations -Réaliser le dashboard de suivi des locations. -Mettre le suivi du leasing avec le dashboard de location en suivant le même format - Ajouter à ce dashboard un onglet de suivi global regroupant à la fois les locations et le leasing. Projet SIMplissime concernant le suivi des eSIM: -Créer un référentiel des cartes SIM permettant d'avoir le maximum d'informations sur celle-ci. - Réaliser un dashboard basé sur des diapositives PowerPoint réalisées par le métier. Projet initié par la stagiaire: Création d'un glossaire commun à toute l'équipe et à tous les dashboards. Participer aux tâches hebdomadaires de l'équipe concernant l'alimentation des datamarts SAS et des sources de données Dataviz.

Techniques statistiques : Analyse des données

Logiciels : SAS Base/ Tableau/ Jira/ Confluence

Promotion 2021 :

Lieu du stage : 94110 ARCUEIL

Intitulé du stage : Data Science

Missions(s) : Modélisation de comportements/ Analyse de flux / Reporting / Segmentation / Analyses de durée de vie

Techniques statistiques : Automatisation des données / Modélisation / Machine learning/ Data visualisation

Logiciels : Dataiku, SAS, SQL, Python, R ou de SPSS.

Promotion 2018 :

Lieu du stage : 78 rue Olivier de Serres Paris

Intitulé du stage : Analyse de données marketing (vente, achat, stock, prévision de ventes) des pays Orange (zone Europe et Middle-East/Afrique)

Missions(s) :

- Analyse de données marketing (vente, achat, stock, prévision de ventes) des pays Orange (zone Europe et Middle-east/Afrique) : def et fiabilisation du processus de récupération de données (format MS ACCESS), mise en place requetes SQL d'intégration, analyse avec les utilisateurs cibles. Mise en place des requêtes SQL d'analyse (avec VBA + ThinkCell), propositions d'axes d'analyses marketing.
- Analyse des sources alternatives de données (réseaux sociaux, usages applicatifs etc).

Techniques statistiques :

Logiciels : Access, SQL, VBA, R

Promotion 2017 :

Lieu du stage : 12F Rue du Patis Tatelin, 35700 Rennes

Intitulé du stage : Analyse de données d'usage client

Missions(s) :

- Définir avec le responsable de stage les cas d'études à analyser (par exemple :
- catégorisation clients, évolution du panier moyen, types de comportements d'usages)
/Identifier les sources de données exploitables pour traiter ces cas (fichiers de logs des portails et équipements, entre autres)
- Identifier et déployer des outils propres à explorer, analyser, et traiter ces données
- Analyser ces données pour évaluer leur qualité, les éventuels traitements de préparation à leur apporter (complétude, suppression des valeurs aberrantes, etc.), puis identifier les données d'intérêt pour les cas d'études envisagés
- Identifier, implémenter (ou développer), puis faire tourner des outils d'analyse susceptibles d'apporter des éléments de réponse aux cas d'études identifiés.

Techniques statistiques : Segmentation, analyse prédictive

Logiciels : R

15 Luxe

CHRISTIAN DIOR COUTURE est société anonyme à conseil d'administration, elle est en activité depuis 1946. Elle est située à PARIS et spécialisée dans le secteur d'activité du commerce de détail d'habillement en magasin spécialisé. Son effectif est compris entre 2000 et 4999 salariés. Sur l'année 2019 elle a réalisé un chiffre d'affaires de 1 869 773 000 €. Elle est considérée comme une grande entreprise (GE).

Lieu du stage : 30 avenue Montaigne – 75008 Paris

Promotion 2020

Intitulé du stage : Assistante Data Analyst

Missions(s) :

- Réalisation d'études et d'analyses clients (analyse du comportement d'achat, profils clients, parcours clients, segmentations)
- Gestion opérationnelle des campagnes marketing : recommandation de ciblage, comptages et créations
- Mise en place de tableaux de bord afin de suivre la performance des campagnes (newsletters, mailing,...)

Techniques statistiques :

- compétences analytiques (lecture de résultats, analyse et communication pertinente des enseignements)
- compétences statistiques à l'aide notamment de SAS

Logiciels : SAS, Excel, pack office, Neolane

15.1 Moët Hennessy

mondial du secteur de luxe de ce marché. Il forme avec le groupe de mode et maroquinerie de luxe Luis Vuitton, le groupe LVMH depuis 1987.

Lieu du stage : Paris

Promotion 2019 :

Intitulé du stage :

Missions(s) : Mise en place de requêtes d'extractions de données de sources diverses et construction des rapports et des Dashboards à l'aide des outils de Business. Conception d'applications pour les services de la supply chain et formation des utilisateurs à la maîtrise des applications.

Techniques statistiques : supply chain, business intelligence

Logiciels : Qlik sense. C'est un nouvel outil de Business Intelligence. C'est une application de visualisation de données qui permet de créer facilement des visualisations plus ou moins complexes, flexibles et interactives...COGNOS

15.2

CAROLL INTERNATIONAL

Caroll, marque de l'élégance et de la féminité, elle met en avant son savoir-faire français, en proposant de très belles pièces. Née dans les années 60, elle habille les femmes avec la même exigence de qualité. Une réussite qui ne se dément pas, elle a réinventé une mode chic, accessible, qui parle à toutes les femmes, aujourd'hui elle rayonne avec un prêt-à-porter complet.

Intitulé du stage : Intégrer le département Marketing Clients pour développer la stratégie Data

Mission(s) : Proposer et mettre en place les modèles statistiques (régression ...) ou de data science (apprentissage supervisé et non supervisé ...) pour résoudre les problématiques métier. / Segmentation et analyse comportementale des clients. / Conception et production de rapports récurrents sur l'activité de l'entreprise (TdB Hebdo, mensuels, KPI clients ...). / Contribuer à l'élaboration des ciblage segments clients.

Techniques statistiques : Modélisation / Analyse de données / Segmentation

Logiciels utilisés : RStudio / SQL / Pack Office

Lieu de stage : 75019 Paris

16 Média

Fondée en mars 2020 à Bruxelles par Nicolas Débué et Loïc Adam, Feelin est une startup à la frontière du marketing et de la data science. Son activité est centrée sur l'analyse des réactions et émotions de l'audience afin d'améliorer la performance des vidéos publicitaires. Ces deux psychologues de formation ont eu l'idée d'utiliser les concepts d'intelligence émotionnelles afin d'évaluer l'impact du marketing numérique. En 2020, Feelin remporte le prix IAB MIXX Awards 2020 dans la catégorie des outils marketing numérique. Les IAB Mixx Awards sont une initiative internationale organisée en Belgique par l'Association belge du marketing et Best of Publishing, éditeur de Inside Magazine et Digimedia.be

Lieu de stage :Belgique

Promotion 2021 :

Intitulé du stage:Data science / Exploration des données comportementales, physiologiques et déclaratives des utilisateurs de l'application FEELIN.

Mission(s): Assister le Lead Data Scientist dans le post processing des données physiologiques Assister le Lead Data Scientist dans l'amélioration des modèles interne de machine learning Explorer des nouvelles méthodes pour la fusion multimodale de données (méthode mathématique, apprentissage supervisé.) Analyser les données de gyroscope et d'accéléromètre pour implémenter des fonctions de correction des données sur base de ces indicateurs Créer des outils de visualisation des données brutes des utilisateurs accessibles online pour usage interne de la société Créer des modèles de rapport de données via libraires ad hoc (Markdown, JupiterNotebook)

Techniques statiques:Analyse de données / Prédiction / Benchmarking

Logiciels utilisés :SPSS, AWS, Python, R, MongoDB

Lieu de stage :Belgique

16.1 Havas Média France

Havas Media Group est l'un des plus grands groupes mondiaux de communication. Le Groupe est présent dans tous les secteurs et dans tous les métiers de la communication. Au sein du département Etudes CSA Data Consulting, véritable Dataroom de l'agence dont l'activité est centrée sur l'analyse d'efficacité et les stratégies media marketing, au service de l'opérationnel.

Lieu du stage : 10 Rue Godefroy 92800 Puteaux

Promotion 2017 :

Stage 1 :

Intitulé du stage :

Missions(s) : Mesurer les impacts des campagnes media et marketing de marque sur les ventes en magasin + site internet./Une analyse allant jusqu'à la mesure de l'ensemble des actions marketing avec un focus particulier des effets de chacun des leviers media. /Prévoir l'évaluation du marché publicitaire Français sur l'ensemble des médiats (TV, Radio, Affichage, Digital et Presse).

Techniques statistiques : Clustering

Logiciels : R, SAS

Stage 2 :

Intitulé du stage :

Missions(s) : Mise en place d'un modèle d'attribution sur des données individuelles digitales. Choix des outils en amont du projet (SAS, R, Python. Une partie de veille sur les différentes approches existantes sera nécessaire pour ce projet. /Reporting de deux tableaux de bords /Travail sur les projets de modélisation clients (secteur à définir en fonction des projets en cours parmi les secteurs Automobile, Grande distribution, Telecom...) en utilisant les process SAS existants. La prise en main du process statistique est un rôle clé dans les analyses et les insights media restitués au client.

Techniques statistiques : Régression linéaire, Séries chronologiques

Logiciels : SAS, R, Python

16.2

Médiamétrie

Médiamétrie, leader des audiences médiatiques, mesure ces audiences (télévisuels, internet et radio) et analyse les évolutions et comportements du public. Les études de Médiamétrie vont par exemple pouvoir permettre aux entreprises de diffuser leurs publicités au moment le plus adéquat pour toucher le public visé.

Lieu de stage : 92300 Levallois-Perret

Intitulé du stage : Détection et traitement de valeurs atypiques

Mission(s): - Recherche bibliographique sur les méthodes de détection et de traitement des valeurs aberrantes ou extrêmes / Etude des méthodes retenues appliquées sur les panels Ordinateur, Mobile et Tablette / Implémentation sous SAS des méthodes possibles et une analyse des résultats d'audience obtenus / Restitution des différentes propositions d'analyses à la Direction

Techniques statistiques : Modélisation

Logiciels utilisés : SAS

16.3 Phenix

Phenix est une street media company. Créée en 2012 en tant que startup technologique, Phenix est devenu l'un des leaders français de la communication extérieure et 1er groupe indépendant du secteur, en tant qu'opérateur propriétaire de réseaux média outdoor et de solutions technologiques en France.

Lieu du stage : 92624 Gennevilliers

Promotion 2021 :

Intitulé du stage : Data Scientist

Mission(s): Structurer et modéliser les données pour les rendre facilement accessibles aux interlocuteurs business

- Définir les KPI à suivre et construire des tableaux de bord automatisés pour piloter les performances et aider les équipes à prendre les bonnes décisions stratégiques au quotidien
- Explorer la donnée à la recherche d'insights, formuler des recommandations sur la base des analyses
- Evangéliser les équipes Phenix sur les sujets data
- Aider les phéniciens à monter en compétence sur les sujets data et à être de plus en plus autonome dans leur utilisation de la donnée

Techniques statistiques : Modélisation, Tableaux de bord

Logiciels utilisés : BigQuery, Looker, Cumul.io

16.4 Prisma média

Prisma Media est une entreprise spécialisée dans la presse. Elle est aujourd'hui le leader des groupes bimédias, de la presse magazine, de la vidéo en ligne et de l'audience digitale quotidienne en France.

Lieu du stage : 92624 Gennevilliers

Promotion 2021 :

Intitulé du stage : Data Scientist

Mission(s): Prédiction socio démo de l'audience et scoring / Segmentation Sémantique NLP / Analyse descriptive & insight / Dashboarding & data viz Communication & pilotage et coordination de projets

Techniques statistiques : Scoring / Segmentation / Analyse de données / Tableaux de bord

Logiciels utilisés : Python

Promotion 2018 :

Intitulé du stage :

Missions(s) : Pôle data science : Stock sur Hadoop /Spark, requêtage avec Dataiku, R, Python. Réaliser des scores, enrichissement de la BDD via open data, prédiction d'audience (modele de prediction), segments(avec google analytics), analyse pression commerciale (email). Le stagiaire participera également aux missions big data : branchements Hadoop/Spark/R avec l'équipe IT

Techniques statistiques :

Logiciels : Hadoop, Bigquery, R, Python, Datakiu

17 Recherche

Le bureau de recherches géologiques et minières (BRGM) est un établissement public à caractère industriel et commercial (Epic) placé sous la tutelle du Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche (MENESR), du Ministère de la Transition Ecologique et Solidaire (MTES) et du Ministère de l'Économie, de l'Industrie et du Numérique (MEIN). Il est l'établissement public de référence dans les applications des sciences de la Terre, pour gérer les ressources et ICS risques du sol et du sous-sol. Le BRGM emploie plus de 1000 personnes dont plus de 700 chercheurs et ingénieurs (géologues, géotechniciens, hydrogéologues, géochimistes, modélisateurs, géophysiciens, informaticiens, ...), dans ses 29 implantations en France métropolitaine et en Outre-mer. Ses équipes interviennent dans une trentaine de pays.

Lieu du stage : 3 Avenue Claude Guillemin, 45100 Orléans

Promotion 2017 :

Intitulé du stage : Utilisation des techniques du Big Data pour rendre accessible les données de qualité de l'eau au robinet

Missions(s) : Le stage s'est déroulé suivant les étapes ci-après : /Etape 1 : Analyse des données de qualité de l'eau potable, de leur variabilité. /Etape 2 : Analyse des solutions pour aspirer ces données, et développement de la solution technique permettant de les récupérer. /Etape 3 : Définition des blocs hadoop de transformation, d'enrichissement et d'analyse, pour déploiement de ceux-ci dans la plateforme hubeau. /Etape 4 : Co-construction des paramètres de l'API, en s'appuyant sur les ré-utilisateurs connus sur les données eau potable. /Etape 5 : Test de l'API, une fois mise en œuvre, et analyse des données correspondantes. /Etape 6 : Créations d'outils d'aide à l'emploi de l'API

Techniques statistiques : HDFS, NoSQL, Webscraping,

Logiciels : Excel, HIVE, JDBC, Hadoop, Spark

Promotion 2018 :

Intitulé du stage : Prédiction du phénomène sismique

Missions(s) :

- Initier une base de données d'apprentissage robuste, réunissant des enregistrements sismiques (larges bandes et/ou accélérométriques), accompagnés de nombreuses métadonnées, issus de différentes bases de données internationales (e.g. K-NET and KiK-net, PEER, ESM, etc.).
- Réaliser le code qui permet de transformer ces ensembles de données en un problème abordable par des méthodes de data science (feature engineering), notamment dans un cadre d'apprentissage automatique supervisé. Tester plusieurs approches d'apprentissage artificiel (e.g "random forest", "xgboost", "SVM", "deep learning"...) sur certaines statistiques du signal sismique, établies sur une fenêtre de temps glissante, comme caractéristiques descriptives. La sélection des données d'entrée, des caractéristiques descriptives, de la taille de la fenêtre de temps et des méta-paramètres des algorithmes de machine learning sont autant de variables du problème à déterminer.
- Réaliser le code permettant la comparaison minutieuse des résultats à travers une campagne de tests systématiques. Les données étant potentiellement très volumineuses, le test intensif et systématique de nombreux modèles engendrera probablement un problème de temps de calcul. Le cas échéant, nous projetons de prendre en compte cette difficulté en intégrant et adaptant les algorithmes d'apprentissage à un cadre de calculs distribués, comme celui proposé par l'environnement Big Data du BRGM

Techniques statistiques : Machine learning Apprentissage supervisé

Logiciels :

17.1 Centre de Calcul Régional ROMEO (Université de Reims Champagne Ardenne)

Le Centre de Calcul Régional ROMEO est une plateforme technologique de l'Université de Reims Champagne-Ardenne depuis 2002. Sa mission est de fournir aux industriels et chercheurs de la région des ressources de calcul performantes, des espaces de stockage sécurisés, des logiciels adaptés, un accompagnement dans l'utilisation de ces outils ainsi qu'une expertise sur des domaines scientifiques et techniques avancé

Lieu de stage:51100 Reims

Promotion 2021 :

Stage 1 :

Intitulé du stage:

Mission(s): Enedis et le Grand-Reims ont mis en place une plateforme de supervision de la consommation d'énergie des 270 points de livraisons de la ville de Reims, sous forme de tableau de bord que les experts peuvent interroger. Il faut installer l'outil dans l'infrastructure de

Cloud ROMEO et de l'adapter pour permettre son passage à l'échelle. Si l'on vise dans un premier temps l'ensemble des points de livraison de la ville de Reims (400), on peut rapidement se projeter sur le périmètre du Grand Reims (143 communes) et d'autres communautés urbaines. In-fine, on souhaite créer une plateforme ouverte opérée par ROMEO, où un client (le GrandReims), pourra donner accès à ses données de consommation.

Techniques statiques:Modélisation / Deeplearning

Logiciels utilisés:

Stage 2 :

Intitulé du stage : Data Science, Machine Learning, Deep Learning

Mission(s) : Développement d'un use-case sur les données de consommation de la ville de Reims : détection de l'occupation des bâtiments et d'anomalies par méthodes de clustering (K-means, MLlib, ..) / Définition d'une API pour la création d'une plateforme pour permettre l'accès aux données et la mise au point de modèles Deep Learning-Machine Learning dans des conditions de production. / Expérimentation de cette API par les use-case identifiés.

Techniques statistiques : Analyse de données / Deep learning / Tableaux de bord

Logiciels utilisés : Python, Excel, MongoDB, NoSQL

Lieu de stage : 51100 Reims

17.2 Centre de recherche en STIC (CReSTIC) / Binus university

- Le Centre de recherche en STIC est une unité de recherche de l'Université de Reims Champagne-Ardenne, issue de la fusion en 2004 du LAM et du LERI. Celui-ci est composé de deux départements, "informatique" et "traitement automatique du signal", et de neuf équipes de recherche. Le CReSTIC se situe sur 4 sites, qui sont Reims, Troyes, Châlons-en-Champagne, et Charleville-Mézières. L'unité est composée d'environ 120 personnes, dont plus de 70 enseignants chercheurs.
- Bina Nusantara, fondée en 1996 par Mr. Joseph Wibowo Hadipoespito et Dr. Theresia Widia Soerjaningsih, est un organisme d'éducation privé situé en Indonésie. Cet organisme est décomposé en trois branches principales :
 - Binus School, pour les jeunes étudiants non-diplômés
 - Binus Higher Education Systems, contenant toute la partie des étudiants diplômés, allant de la licence, jusqu'au doctorat.
 - Binus Enterprises, comportant les organismes hors académique, tel que Binus Square, qui est une résidence étudiante.

Lieu du stage : Champs-sur-Marne

Promotion 2022 :

Intitulé du stage : Sarcasm detection in Twitter

Missions(s) : Utiliser des données de type texte afin de détecter si une phrase est sarcastique ou non.

Techniques statistiques : Deep learning

Logiciels : python

17.3

CNRS Délégation Centre-Est

L'URCA est composée de plusieurs unités de formation et de recherche (UFR) comme l'UFR de Médecine, l'UFR de Sciences économiques, l'UFR de Sociales et de Gestion, l'UFR de Sciences exactes et naturelles (SEN) et bien d'autres. L'UFR SEN comprend plusieurs disciplines dont celles de l'informatique, de la physique, de la biologie ou encore des mathématiques. Ces disciplines ont chacune un ou plusieurs laboratoires affiliés ; par exemple, l'UFR SEN comprend 11 laboratoires de recherches, dont celui de mathématiques.

Lieu de stage : 54519 VANDOEUVRE

Promotion 2021 :

Intitulé du stage : Modélisation et prévisions des flux de patients dans les services d'urgences

Mission(s) : L'objectif est de développer des outils statistiques de prévisions des flux de patients dans les Services Administratifs d'Urgence (SAU) de la région Grand Est. À partir d'un historique de données du nombre d'arrivées de patients dans les SU, la/le stagiaire devra mener une analyse statistique complète de ces données allant du prétraitement à la construction de modèles de séries temporelles prédictifs

Techniques statistiques : Analyse de données / Modélisation / Prédiction / Tableaux de bord

Logiciels utilisés : RStudio

17.4 Institut de Recherche bio-Médicale et d'Épidémiologie du Sport

Depuis 1975 l'INSEP existe tel qu'il est aujourd'hui, ayant pour mission de participer « à la recherche scientifique fondamentale et appliquée en matière pédagogique, médicale et technique; à la formation continue de niveau supérieur des personnels enseignants d'éducation physique et sportive, des conseillers techniques et des éducateurs sportifs ainsi que des personnels des services de la jeunesse et des sports; à l'entraînement des équipes nationales ainsi qu'à la promotion des sportifs de haut niveau »

Lieu du stage : 11 avenue de Tremblay – 75012 Paris

Promotion 2022 :

Intitulé du stage : ESTIMATION DU RISQUE DE BLESSURE ET CRÉATION D'UN INDICATEUR DE DENSITÉ DES CHARGES D'ENTRAÎNEMENT AU RUGBY

Missions(s) : Préparation des jeux de données issus de différentes sources (Questionnaire du matin, données de tests physique, données GPS d'entraînement et de matchs, blessures...), l'exploration et statistique descriptive, la création d'un indicateur de densité d'entraînement à partir des données GPS, la mise en place de modèles d'estimation de risques de blessures et la rédaction d'un article scientifique inhérent aux tâches précédentes

Techniques statistiques : Visualisation des données

Logiciels : R

17.5 Laboratoire de Mathématiques de Reims

Le Laboratoire de Mathématiques de Reims (LMR), créé en 1992, est aujourd'hui dirigé par Michael Pevzner. Les thématiques de recherches abordées sont très larges, « allant de la théorie des nombres et théorie des représentations à la modélisation numérique et stochastique en passant par l'analyse des EDP, la théorie spectrale et les probabilités ». Le laboratoire se compose de 3 équipes de recherche : « Analyse », « Groupes et Quantification » et enfin « Modélisation Numérique et Stochastique », dont je fais partie. Au total on compte 49 personnes appartenant au laboratoire, ce qui comprend entre autres les enseignants-chercheurs, les doctorants, les ingénieurs de recherche et le personnel administratif.

Lieu du stage : Champs-sur-Marne

Promotion 2022 :

Intitulé du stage : analyser les données volumétriques cérébrales des nouveau-nés prématurés

Missions(s) : Analyser les données volumétriques cérébrales des nouveau-nés prématurés. Mettre en place un score et une représentation graphique simple permettant d'aider à la validation des segmentations par les cliniciens.

Techniques statistiques : Modélisation statistique / Analyse des données

Logiciels : R/Latex

17.6 Lab'URBA Université Paris-Est

Le Lab'URBA se définit principalement par son inscription dans l'Ecole d'Urbanisme de Paris issue de la fusion des deux principaux instituts d'urbanisme français.

Le périmètre du Lab'URBA associe également des enseignants-chercheurs partageant ses valeurs et travaillant sur les questions urbaines, notamment au sein du département de géographie de l'Université Paris Est Créteil (UPEC), du département Génie Urbain de l'Université Gustave Eiffel (anciennement Université Paris-Est Marne-la-Vallée) et de l'Ecole des ingénieurs de la Ville de Paris (EIVP).

Ces instituts et départements trouvent leur raison d'être dans une forte association entre une formation qui suit les évolutions de ses mondes professionnelles de référence, en France et dans le monde, et contribue à cette évolution par la production de connaissances, la mise en œuvre de ces connaissances dans des activités d'expertise et un rôle d'animation dans le débat public sur les choix sociaux qui concernent la production, la gestion et l'usage des villes.

Lieu du stage : Champs-sur-Marne

Promotion 2019 :

Intitulé du stage : réalisation d'une enquête visant à identifier une typologie des touristes susceptibles de venir aux JO et Paralympiques de Paris 2024

Missions(s) : Réaliser une enquête visant à identifier les pratiques et la mobilité des touristes qui viennent à Paris et en Ile-de-France

Techniques statistiques : Echantillonnage, analyse de données

Logiciels : R

17.7 NEOMA Business School

NEOMA BS est une école qui propose des formations dans l'ensemble des disciplines de sciences de gestion: marketing, gestion, finance, droit, commerce. Elle offre des formations allant du bac+3 au bac+5 dans le domaine du management, du commerce et de l'entreprise et met la recherche au centre de ses activités afin de sans cesse innover et s'adapter à toutes les problématiques nouvelles.

Lieu du stage : Reims

Promotion 2019 :

Intitulé du stage : Panorama technologique des micro-algues: identification des acteurs du champ technologique des micro-algues et leurs positionnements

Missions(s) : Identification des acteurs scientifiques et technologiques du domaine des micro-algues respectivement à partir de l'activité de publication et de l'activité de dépôt de brevets

Techniques statistiques : Bibliométrie, Analyse de systèmes complexes

Logiciels : SQL, R, Gephi

17.8 Université de Paris, MAP5 (CNRS UMR 8145)

Le MAP5 est rattaché à Université de Paris (ex Université Paris Descartes) et à l'Institut National des Sciences Mathématiques et de leurs Interactions (INSMI) du CNRS. Il fait partie de la Fondation Sciences Mathématiques de Paris (FSMP).

Le MAP5 est un laboratoire récent créé sous sa forme présente en 2004. Il a pris la suite d'équipes principalement tournées vers les applications de la statistique au domaine biomédical et vers l'image. Son effectif est passé de 28 permanents en 2005 à une cinquantaine aujourd'hui. Une des ambitions du laboratoire est d'établir à Université de Paris un excellent niveau de recherche en mathématiques appliquées, tout en développant des liens avec les autres composantes de l'université. Le MAP5 est composé de 4 équipes, travaillant chacune dans un domaine particulier des mathématiques appliquées : les statistiques, les probabilités, l'analyse numérique et le traitement des images. Il rassemble une variété d'expertises et de compétences assez large pour créer une vraie synergie entre les développements théoriques et les applications.

Lieu du stage : 85 Bd Saint Germain, 75006 Paris

Promotion 2010 :

Intitulé du stage :

Missions(s) : modelisation, developpement , entrainement et implementation d'un reseaux de neurones

Techniques statistiques : modélisation statistique, machine learning

Logiciels : R, matlab, python, tensorflow, keras

17.9 UPEM (Université Paris-Est-Marne-la-vallée)

La multiplicité des phénomènes urbains pose de nouveaux défis qui ne peuvent être résolus qu'en appliquant une approche pluridisciplinaire. Des moyens importants sont mis en œuvre pour donner une envergure internationale aux projets de recherche des laboratoires compétents dans ce domaine. Le LabEx (Laboratoire d'Excellence) Futurs Urbains a été créé en 2011, suite à un appel d'offre des investissements d'avenir ; 3000 projets pour une enveloppe de 47 Milliards d'Euros. L'idée étant de mettre en commun les compétences de chercheurs issus de milieu professionnels différents, tels que l'urbanisme, l'architecture, l'environnement, l'énergie, etc. On est alors en mesure d'obtenir une vision plus précise des interactions en environnement urbain ; des gens entre eux, ou avec les infrastructures par exemple, ou encore l'impact de l'interdiction des vieilles voitures en centre-ville.

Lieu du stage : 5 Boulevard Descartes, 77420 Champs-sur-Marne

Promotion 2017 :

Intitulé du stage : Mise en perspective des données d'enquête et des données issues des terminaux mobiles : cas du marché de noel des Champs-Élysées 2016-2016

Missions(s) :

- Réaliser une étude comparative des données
- Réaliser une carte graphique permettant de visualiser la provenance des touristes français et étrangers
- Compléter l'analyse qui a été produite l'année précédente

Techniques statistiques : Data visualisation, Régression linéaire

Logiciels : R, Excel

18 Services de la santé

Lieu de stage : 10003 Troyes

Promotion 2021 :

Intitulé du stage : La gestion des tensions hospitalières par l'intelligence artificielle

Mission(s) : Se familiariser avec les données hospitalières / Mettre en place un dashboard pour la visualisation de données - Extraire de données utiles (sélection, extraction, nettoyage, etc) / Élaborer un modèle pour la quantification et la modélisation de la tension / Développer un algorithme intelligent pour l'anticipation de la tension / Développer un système d'aide à la décision pour la gestion des épisodes de tension.

Techniques statistiques : Prédiction / Modélisation / Tableaux de bord

Logiciels utilisés : RStudio / Excel / Power BI

18.1 Cosmetic Dermatology LLC (DR BRANT)

Frederic Brandt, est un médecin chirurgien spécialisé en dermatologie. Il est considéré comme étant un spécialiste de la peau. Il décide en 1995, de lancer sa marque de cosmétique. Il décide donc de fonder DrBrandt Skincare. Ainsi les produits, qu'il propose sont des produits innovants. En effet, Les produits de la gamme Dr Brandt sont conçus afin de respecter 3 critères importants. Ces critères sont :

- Les produits sont basés sur la recherche scientifique et dermatologique
- Les produits sont efficaces
- Les produits sont sans danger pour tous types de peau

Dr Brandt fut l'un des premiers à s'intéresser aux questions relatives au soin de la peau. A titre d'exemple, il fut l'un des premiers à commercialiser des produits pour lutter contre l'acné.

Lieu de stage : 75008 PARIS

Promotion 2021 :

Intitulé du stage : Data Analyst

Mission(s) :

- Conception et mise en place d'une solution de web tracking
- Collecte des données et analyses
- Optimisation du site web Shopify pour l'UX et les moteurs de recherche (on-page SEO)
- Evaluation de l'expérience utilisateur du site et des taux de conversion
- Mise en place des recommandations SEO.

Techniques statiques : Webscrapping / Prédiction / Tableaux de bord

Logiciels utilisés : Salesforces / Shopify

18.2 KAP CODE

Kap Code est une start-up spécialisée dans le développement de solutions e-santé innovantes à destination des patients atteints de pathologies chroniques.

Lieu du stage : 28 rue d'Enghein, 75010 Paris

Promotion 2017 :

Intitulé du stage : Détection et analyse de différents types de mésusages de Paracétamol sur des données textuelles provenant des réseaux sociaux dévolus à la santé.

Missions(s) : L'objectif a été de répondre à une demande de Sanofi (qui est un groupe pharmaceutique et client de Kap code) qui est de pouvoir détecter la présence de trois cas de mésusage des médicaments composés d'au moins du paracétamol. Cette étude permettra à Sonofi d'avoir un suivi dans le temps de l'opinion des patients en amont et en aval de la mesure de prévention sur le surdosage du paracétamol.

Techniques statistiques : Extraction des données (collecte des données et la mise en forme du dataset), le prétraitement des données (filtrage des données, traitement des valeurs manquantes, dédoublonnage)

Logiciels : Rstudio, Excel

18.3 Logever Healthcare

Logever-Healthcare, lancé en 2008, est l'entité qui est chargé de l'édition, développement et commercialisation de Logicoss. Logicoss a été pensé dans le but de simplifier et rendre fiable les opérations de 8 missions clés, à savoir : - Suivre les patients en poste de secours / - Prendre une alerte au poste de Commande lorsqu'un participant se blesse au cours de l'évènement. / - Engager une équipe de secours pour prendre en charge le patient. • Diriger et superviser les secours. / - Accélérer la remontée d'alerte. / - Fournir des outils de traçabilité aux organisateurs et permettant aussi de protéger de façon juridiques le personnel soignant. / - Informer les autorités en cas de risques majeurs. / - Informer les familles sur l'état et l'emplacement du participant, toujours en course ou transféré dans un hôpital.

Lieu de stage : 69310 Pierre Benite

Promotion 2021 :

Intitulé du stage : Sport Events

Mission(s) : Exploitation de la base constituée des données de prises en charge médicales de 500 000 sportifs et de 10 millions de visiteurs de sites tertiaires par an. / Poursuivre les travaux du modèle prédictif d'accidentologie débuté en 2019. / Fournir les premiers outils d'analyse vers l'optimisation des plannings d'ambulance dans le département du Rhône.

Techniques statistiques : Analyse de données / tableaux de bord

Logiciels utilisés : RStudio / Python / Gitlab

19 Voyage & Tourisme

TUI (Travel Union International) est une société anonyme en activité depuis 1984. Elle évolue dans le secteur des activités des voyagistes. Son effectif est compris entre 1000 et 1999 salariés faisant une entreprise de taille intermédiaire.

Lieu du stage : 32 rue Jacques Ibert|92309 Levallois-Perret Cedex|France

Promotion 2019 :

Intitulé du stage : Segmentation clients et analyse de sentiments

Missions(s) : Réaliser une étude visant à identifier la catégorie d'un client et de le classer suivant son comportement d'achat par la segmentation selon la ressemblance et l'analyse de sentiment après consommation du produit

Techniques statistiques : Segmentation (méthode K-means) et analyses de sentiments (text mining: c'est une méthode qui permet de tirer de l'information utile à partir des commentaires laissés par les clients après consommation du produit. Cela permet d'apprendre sur les points forts et les points d'améliorations de l'entreprise et donc d'obtenir une vue d'ensemble)

Logiciels : SQL, Python et Power BI

20 Autres

Alchimie Dynamic Digital Distribution une société par actions simplifiée qui exerce dans la région parisienne à Aubervilliers (93300) depuis 6 ans. Son domaine d'activité est le secteur des activités des sociétés holding. Elle appartient à la catégorie des entreprises de taille intermédiaire.

Lieu du stage : 45 Avenue Victor Hugo, 93300 Aubervilliers

Promotion 2019 :

Intitulé du stage : Elaboration de tableaux de bord et d'analyses facilitant la gestion et la prise de décision

Missions(s) : Réaliser une transformation des données ainsi que des événements bruts de façon plus visuelle pour que cela puisse être compris par les autres services notamment le service marketing.

Techniques statistiques : Web analytics

Logiciels : Splunk (c'est un logiciel d'analyse de données via une interface web. Il permet notamment de comprendre et d'analyser les parcours des utilisateurs sur un site web.), SQL et Excel.

20.1 CENTRE VINICOLE CHAMPAGNE NICOLAS FEUILLATTE (CV-CNF)

Aujourd'hui, le CV-CNF représente la première union de producteurs de Champagne. Il compte plus de 80 coopératives et une Union des Producteurs Particuliers (UPP-CVCNF) composée de plus 1360 adhérents. Il compte au total près de 5000 vignerons. Sa mission principale 5 aujourd'hui est d'apporter le savoir-faire et une qualité à l'élaboration des vins de Champagne qui assurent une rémunération optimale des apports à ses adhérents et une réputation mondiale.

Lieu de stage : 51530 CHOUILLY

Promotion 2021 :

Intitulé du stage : Stage gestion et optimisation Data au sein du service des relation vignoble du CV-CNF

Mission(s) : Automatisation de rapports data vendanges (avant, pendant, après) / Refonte des rapports de pilotage du service et création de nouveaux rapports

Techniques statistiques : Analyse de données / Cartographie / Tableaux de bords

Logiciels utilisés : JDEdwards / Cognos

20.2 Daimler AG

Daimler Mobility AG se focalise sur les solutions financières et de mobilité visant à faciliter la vie quotidienne de ses clients. Les services vont des offres de location, de financement aux services de gestion de flotte, aux assurances et plateformes de paiement électronique. DMO investit également dans des solutions de mobilité basées sur des applications pour covoiturage, la mobilité en tant que service (MaaS) et autres.

Lieu de stage : 70320 Stuttgart-Allemagne

Promotion 2021 :

Intitulé du stage : Data science

Mission(s) :

- Travailler avec les équipes d'experts sur l'amélioration des modèles de classification quantitative des risques, (par exemple le scoring) et d'évaluation avancées utilisant des méthodes d'apprentissage automatique
- Préparation des données, analyse des données ainsi qu'interprétation et présentation des résultats
- Participation à des projets internationaux mettant en œuvre des processus et des solutions de décision en matière de crédit numérique
- Évaluation de l'applicabilité des méthodes avancées du machine learning dans la classification des risques.

Techniques statistiques : Analyse de données / Scoring/ Classification

Logiciels utilisés : RStudio / Python / SAS / Spark

20.3 Pari Mutuel Urbain (PMU)

Le PARI MUTUEL URBAIN est un groupement d'intérêt économique créé en 1930. Installée à Paris, elle est spécialisée dans le secteur d'activité de l'organisation de jeux de hasard et d'argent. Son effectif est compris entre 1000 et 1999 salariés. Sur 2019 elle a réalisé un chiffre d'affaires de 9561 milliards d'euros. Elle appartient à la catégorie entreprise de taille intermédiaire (ETI).

Lieu du stage : 2 Rue du Professeur Florian Delbarre 75015, Paris

Promotion 2017 :

Intitulé du stage : Optimisation de la montée en gamme des clients

Missions(s) :

- Analyse d'opération marketing et statistiques descriptives;
- Modélisation dans le but de déterminer les facteurs les plus importants pour un up sell avec la plus forte valeur ajoutée;
- Test d'opération marketing pour valider l'analyse et la modélisation en vue

Techniques statistiques : Segmentation, K-mens, Régression logistique

Logiciels : SAS, Excel, Powerpoint

20.4