



ÉCOLE NATIONALE
DES SCIENCES
GÉOGRAPHIQUES

École Nationale des
Sciences Géographiques



Galigeo

Rapport de stage

Cycle des Ingénieurs diplômés de l'ENSG 3^{ème} année

Stage de fin d'étude ENSG

Estimation, analyse et prédiction de flux piétons

BigData et Machine Learning



ÉCOLE NATIONALE
DES SCIENCES
GÉOGRAPHIQUES

Jules Pierrat

Septembre 2022

☒ Non confidentiel ☐ Confidentiel IGN ☐ Confidentiel Industrie ☐ Jusqu'au ...

ÉCOLE NATIONALE DES SCIENCES GÉOGRAPHIQUES
6-8 Avenue Blaise Pascal - Cité Descartes - 77420 Champs-sur-Marne
Téléphone 01 64 15 31 00 Télécopie 01 64 15 31 07

Jury

Président de jury :

Victor Coindet Professeur de l'ENSG, Responsable du cycle TSI

Commanditaire :

M. Sebastien Connesson, COO de Galigeo

Encadrement de stage :

M. Jean-Michel Gaudin, Responsable du pôle Recherche et Développement à Galigéo

Enseignant référent :

M. Loic Landrieu, Chercheur MATIS / IGN, Professeur de l'ENSG

Rapporteur expert :

qui est rapporteur du mémoire ?

Responsable pédagogique du cycle Ingénieur - TSI :

Victor Coindet Professeur de l'ENSG, Responsable du cycle TSI

Gestion du stage :

Delphine Genès, Relation entreprise de l'ENSG

Stage de fin d'étude du 2 mai 2022 au 28 Octobre 2022

Diffusion web : ☒ Internet ☒ Intranet Polytechnicum ☒ Intranet ENSG

Situation du document :

Rapport de stage de fin d'études présenté en fin de 3^{ème} année du cycle des Ingénieurs

Nombres de pages : 25 pages dont 5 d'annexes

Système hôte : L^AT_EX

Modifications :

EDITION	REVISION	DATE	PAGES MODIFIEES
1	0	09/2022	Création

Remerciements

Avant toute chose, je tiens à remercier le lecteur pour l'intérêt qu'il porte à mon rapport et j'espère qu'il trouvera ici tout ce pourquoi il est venu. Je veux remercier également les personnes m'ayant permis de réaliser dans les meilleures conditions ce stage ainsi que celles ayant contribué à l'élaboration de ce rapport.

Tout d'abord, j'adresse mes remerciements à mon professeur, **Mr Loïc Landrieux de l'Ecole Nationale des Sciences Géographiques**, mon maître de stage qui m'a suivi tout au long de ce stage, m'a guidé et éclairé dans mes décisions.

Je tiens à remercier mon maître de stage, **Jean-Michel Gaudin, Responsable du pôle recherche et développement à Galigeo** pour son suivi et l'intérêt qu'il a porté à mes travaux réalisés pendant le stage. Je remercie également **Raimana Teina, Data Scientist chez Galigeo** qui m'a guidé dans mes travaux et grâce à qui j'ai énormément progressé et appris durant toute ma période de stage. Mes remerciements vont également à **Sébastien Connesson, COO de Galigeo**, qui m'a permis de comprendre au mieux l'organisation de l'entreprise, les relations internes, les enjeux et les rapports aux clients.

Je remercie évidemment tout le reste de **l'équipe de Galigeo** pour son accueil, la confiance qu'ils m'ont accordée, leurs conseils et leur bienveillance. Je suis très heureux d'avoir pu travailler avec eux et me réjouis de continuer à le faire.

Enfin, après ces trois années fabuleuses je tiens à remercier toutes les personnes qui ont croisé mon chemin à **L'Ecole Nationale des Sciences Géographiques, mes professeurs, mes amis** et toutes les rencontres qui m'ont permis de grandir et de me préparer à cette nouvelle vie après les études.

Enfin, je tiens à remercier toutes les personnes qui m'ont conseillé et relu lors de la rédaction de ce rapport de stage : **ma famille, mon ami Antoine Rainaud** camarade de promotion.

Résumé

Ceci est mon résumé bla bla bla

Mots clés : clés, clés, clés

Résumé

This is my abstract blah blah blah...

Key words : key, key, key

Table des matières

Remerciements	3
Glossaire et sigles utiles	3
Introduction	4
1 Galigeo Ensg	5
1.1 Présentation de l'entreprise	5
1.2 Les objectifs de Galigeo	7
1.3 Organisation du stage	8
2 Le stage Ensg	9
2.1 Généralités	9
2.2 Prédiction de flux piéton	9
2.3 Prédiction de chiffre d'affaire	9
2.4 Autres missions	9
3 Bilan Ensg	11
3.1	11
Conclusion	12
A Filtre de Kalman	23
B Moindres carrés	25

Glossaire et sigles utiles

BI Business Intelligence
CNN Convolutional Neural Network
CEO Chief Executive Officer
COO Chief Operating Officer
CRM Customer Relationship Management
DNN Deep Neural Network
ENSG École Nationale des Sciences Géographiques
LSTM Long Short Term Memory
ML Machine Learning
RNN Recurrent Neural Network
SVM Support Vector Machine

Introduction

Le géomarketing ou Location Business Intelligence en anglais est un pilier du marketing. Il étudie la variation des marchés dans l'espace. Les objectifs sont de modéliser offres et demandes en fonction de données économiques, sociales, culturelles, administratives et démographiques et leurs variations en fonction des géographies.

C'est un domaine essentiel pour les entreprises qui cherche à développer leurs espaces d'action. En effet, c'est une solution qui aide à la prise de décision pour le développement d'un business. Il permet de choisir les sites stratégiques les plus appropriés pour implanter un nouveau commerce. La réalisation de modèle ou de simulation sont des outils essentiels en vue de comparer les atouts et risques d'une future implantation. En étudiant les espaces entourant toutes ses enseignes, une entreprise peut également anticiper la cannibalisation¹ ou la segmentation des portefeuilles² en prenant tout les paramètres réunis en un point de l'espace.

Il permet également d'établir des stratégies de marketing rentables et efficaces en établissant des profils de susceptibles consommateurs. En modélisant de manière précise et orienté dans le sens des besoins de l'entreprise ces profils, on obtient alors une idée complète des déplacements et des comportements réels des consommateurs. Les stratégies de prospection et communication sont donc amené à être plus efficace.

La concurrence est également bien étudiée et cela permet de projeter la pérennité de l'entreprise dans le temps en alertant sur l'évolution des réseaux de concurrents.

Le géomarketing est la solution efficace afin d'appréhender parfaitement les territoires impactés par une nouvelle implantation et ainsi suivre son évolution tout au long de sa croissance.

Ces dernières années, les solutions de géomarketing s'appuie de plus en plus sur des modèles de prédictions de plus en plus complexe et précis. La mise à disposition de modèle de flux piéton s'avère très utile pour permettre d'améliorer le géomarketing d'une compagnie. La demande concernant cette mesure est en augmentation et les entreprises vendeuses de solutions de géomarketing cherche à obtenir les meilleurs modèles prédictifs pour répondre au mieux aux besoins de leurs clients.

Les algorithmes de Machine Learning sont des outils puissants pour estimer spatialement les flux piétons utiles aux analyses de géomarketing. Ils nécessitent cependant des quantités de données très importantes pour obtenir les précisions nécessaires.

Mon stage a consisté en partie à la réalisation de ce modèle prédictif. Ce genre de mission est typique au métier de Géo Data Scientist³ et c'est donc autour de cette mission que je développe mon rapport de stage.

1. A compléter

2. A compléter

3. A completer

1.1 Présentation de l'entreprise

1.1.1 Généralités

Galigéo est une société parisienne spécialisée dans le géomarketing. Créé en 2001, elle propose aux entreprises d'améliorer l'efficacité de tous leurs métiers grâce à ses logiciels combinant expertise cartographique et modélisation prédictive. Grâce à ses logiciels visualisant, analysant et agissant directement sur les bases de données opérationnelles (applications métier, BI, CRM, ...). Galigéo permet aux utilisateurs de se focaliser sur leur métier (retail¹, distribution, marketing, sécurité, ...).

Historiquement pionnier de la location Intelligence, Galigéo a poursuivi son développement ses dernières années en ajoutant la composante prédictive dans ses suites logiciels, composante basée sur les techniques innovantes de Machine Learning et d'Intelligence Artificielle.

En utilisant son expertise cartographique et de modélisation prédictive, Galigéo poursuit son développement en mettant à disposition de ses clients, des logiciels simples d'usage, à très forte valeur ajoutée métier.

De 2001 à 2006, Galigéo se consacre aux développements de solutions intelligentes en géodécisionnel pour faire de l'analytique avancée à partir de cartes géographiques. De 2006 à 2011, elle développe sa première solution logiciel, Galigéo Enterprise, qui permet par la suite, grâce à un nouveau pôle Conseil, de proposer des services spécifiques métier et adapté à chaque client autour du géomarketing. A partir de cette date, Galigéo va voir sa croissance augmenter sur le marché international en s'associant à différents partenaires. A partir de 2017 et jusqu'à aujourd'hui, Galigéo a choisi d'ajouter à ses solutions une composante prédictive très demandé sur le marché.

L'entreprise compte aujourd'hui une cinquantaine de clients très divers, grandes enseignes commerciales, services publics, industries, etc. Galigéo peut fournir des solutions logiciels standards pour permettre à n'importe quelle entreprise de générer des rapports de géomarketing en utilisant des données interne et des données générales fournies par Galigéo. Elle réalise également des projet plus spécifique, propre à des besoins bien déterminés qui permette alors à ces client d'avoir une vraie valeur ajoutée et un géomarketing efficace.



FIGURE 1.1 – Exemple de clients Galigéo

Les bureaux de l'entreprise sont situés au 87 avenue d'Italie dans le treizième arrondissement de Paris.

1. A compléter



FIGURE 1.2 – Bureaux de Galigeo, 87 avenue d'Italie, 75013 Paris

1.1.2 Organisation interne

Galigéo est composé actuellement d'une vingtaine d'employés réparti dans 5 pôles : Administratif, Commercial, Marketing, Comptabilité, Consulting et Recherche et Développement.

Le pôle administratif est le pôle qui s'occupe de la gestion du budget, du personnel et des missions à Galigeo.

Le pôle commercial gère les relations clients, il propose des offres à de nouveaux ou ancien client, prospecte et cherche à faire grandir le cercle de clientèle de Galigeo.

Le pôle marketing imagine les produits, met en place les stratégies de pénétration du marché et réalise un catalogue de solutions que Galigeo peut fournir.

Le pôle comptabilité est responsable de la facturation des clients et la gestion interne des frais de personnels, des salaires, du matériel, etc.

Le pôle consulting est dédié à la réponse aux besoins techniques du client, il permet de rester proche du client. Il imagine les solutions retenues et les intègres dans les outils Galigeo mis à disposition pour l'entreprise.

Pour ma part, j'ai rejoint le pôle Recherche et Développement. L'équipe est composée de développeur, de testeurs, de designer, de data scientist, etc. Elle s'attache à améliorer les produits de Galigeo et à faire du support, de la maintenance et de l'innovation. Lorsqu'un consultant est en charge d'un projet pour un client, il s'appuie sur un ou plusieurs membres de l'équipe R&D pour conseiller ou réaliser les tâches techniques.

J'ai principalement travaillé avec Raimana Teina, Data Scientist chez Galigeo autour du grand projet actuel chez Galigeo « Prédiction de flux piéton » que je détaillerais dans la suite de ce rapport. Cependant j'ai également eu l'occasion de travailler sur des projets clients avec l'équipe consulting.

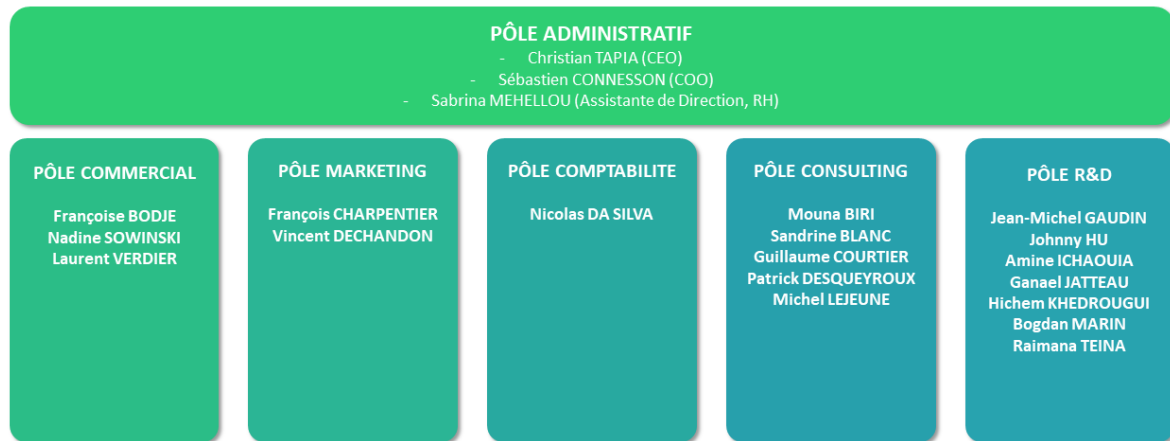


FIGURE 1.3 – Bureaux de Galigeo, 87 avenue d'Italie, 75013 Paris

1.2 Les objectifs de Galigeo

1.2.1 Les manques actuels

Aujourd'hui, Galigeo cherche à mettre en avant des modèles prédictifs au service du géomarketing pour ses clients. Une mesure importante en géomarketing est l'estimation de flux piéton à un endroit donné et sur une période donnée. Jusqu'ici, Galigeo utilisait un service tier afin d'obtenir une estimation de flux piéton. Cette solution de transition possède néanmoins de nombreux défauts. Tout d'abord Galigeo n'avait aucune visibilité sur les algorithmes utilisés pour faire cette estimation.

Il était alors difficile d'utiliser les résultats de cette estimation dans l'entraînement de nouveaux modèles de prédiction (estimation de chiffre d'affaires, estimation de parts de marchés, etc.). De plus, les coûts d'une telle solution restent élevés.

Il a donc été décidé de créer un modèle d'estimation de flux piéton propre à Galigeo sur lequel l'entreprise aurait accès à toutes les données d'entrées du modèle. Ainsi elle pourra modéliser d'autres variables essentielles au géomarketing plus facilement et avec une meilleure qualité. En effet, en connaissant les biais de notre modèle, ils seront plus faciles à corriger dans d'autres cas d'utilisation.

Galigeo souhaite également renforcer son équipe big data afin de mettre en valeur de grosses quantités de données brutes inexploitées en l'état. En recrutant de nouveaux data scientist, elle libérera du temps à ses consultants et déchargera les équipes spécialisées dans la data actuellement en place.

Après un entretien au printemps 2022 chez Galigeo et cette explication des manques et objectifs de l'entreprise à court terme, j'ai fait part de mon envie de réaliser un stage au sein de l'équipe R&D. Ils ont retenu ma candidature et j'ai donc commencé mon stage de fin d'étude le 02 Mai 2022.

1.2.2 Les objectifs du stage

Durant ce stage j'avais donc comme premier objectif de découvrir le géomarketing dans son ensemble. J'ai dû analyser et comprendre les cas d'usages métiers et me plonger dans le fonctionnement d'une équipe de développement logiciel.

J'avais également pour mission de concevoir et spécifier des solutions d'acquisition, de nettoyage et de traitement des données. Comprendre le fonctionnement des produits Galigeo m'a permis de collecter et stocker la donnée de manière optimale afin de faciliter l'implémentation dans des bases de données.

J'avais également pour mission de développer des chaînes de collecte de la donnée en m'appuyant sur des méthodes de data engineering². Pour pouvoir par la suite utiliser au mieux la donnée dans des processus d'analyse et de traitement.

Pour la partie traitement de la donnée brute, mes objectifs étaient de développer des algorithmes et des traitements de la data grâce à des méthodes statistiques et de machine learning. Mon objectif était de transformer de la donnée brute inexploitable en une donnée pertinente pour le géomarketing des clients Galigeo.

Mon objectif était également de m'intégrer aux équipes de Galigeo afin de mieux comprendre les objectifs de l'entreprise à long termes et les directions à prendre.

1.2.3 Les objectifs à plus long termes

1.3 Organisation du stage

1.3.1 Planning

1.3.2 Mes missions

1.3.3 Relations internes et client

2. A compléter

LE STAGE

- 2.1 Généralités**
- 2.2 Prédiction de flux piéton**
- 2.3 Prédiction de chiffre d'affaire**
- 2.4 Autres missions**

3.1 ...

Conclusion

Il est l'heure de conclure : bonne nuit !

Bibliographie

- [1] Balhorn R et AL. "Frequency Stabilization of Internal-Mirror Helium-Neon Lasers". In : *Applied Optics* (1972), p. 742-744.
- [2] Niebauer T.M. et AL. "A new génération of absolute gravimeters". In : *Metrologia* 32 (1995), p. 159-180.
- [3] Niebauer T.M. et AL. "Frequency stability measurements on polarization-stabilized He-Ne lasers". In : *Applied Optics* (1988), p.1285-1289.
- [4] Jacques BEILIN. "Compensation combinées d'observations gravimétriques absolues et relatives – MCGRAVI". Projet de fin d'étude. ENSG, 2005.
- [5] Helmut Moritz BERNHARD HOFMANN-WELLENHOF. *Physical Geodesy*. T. XVII. 2nd, corr. ed., Springer, 2006, p. 403.
- [6] Pilot laboratory BIPM. *7th International comparison of absolute gravimeters, ICAG-2005, Technical protocol*. Rapp. tech. BIPM, 2005.
- [7] BRGM. *Site web du BRGM - SIG Mines France*. BRGM. Nov. 2009. URL : http://sigminesfrance.brgm.fr/geophy_gravi.asp.
- [8] Michel CAPDEROU. *Satellites : orbites et missions*. Springer, 2002, p. 511. ISBN : 9782287597725.
- [9] Frédéric CHAMBAT. "Voir la planète avec la pesanteur". In : *Dossiers Pour la Science* 67 (2010), p. 72-73.
- [10] Michel DIAMENT. "Mesure du champ de pesanteur terrestre". In : *Techniques de l'ingénieur* R 1814 (2005).
- [11] Bernard DUCARME. *MT80 - Theoretical tides computation*. Observatoire Royal de Belgique.
- [12] Françoise DUQUENNE et al. *GPS : localisation et navigation par satellites*. Hermès science publications, 2005, p. 330. ISBN : 9782746210905.
- [13] Henri DUQUENNE. "Altitude, nivellement, systèmes de référence altimétrique". In : *Cours aux étudiants PPMD*. ENSG, 2005.
- [14] Germinal GABALDA et Sylvain BONVALOT. *CG3TOOL, Programme interactif de traitement de données gravimétriques Scintrex CG3/3M*. IRD. Juin 2000.
- [15] Germinal GABALDA, Sylvain BONVALOT et Roger HIPKIN. "CG3TOOL : an interactive computer program to process Scintrex CG3/3M gravity data for high resolution applications". In : *Computer & Geosciences* 29 (2003), p. 155-171.
- [16] Scott GLEASON et Demoz GEBRE-EGZIABHER. *GNSS Applications and Methods*. Artech House, août 2009, p. 528. ISBN : 9781596933293.
- [17] GPSW. *IS-GPS-200 rev e, Navstar GPS Space Segment/Navigation User Interfaces*. Rapp. tech. El Segundo, California : Global Positioning System Wing (Gpsw) Systems Engineering & Integration, 2010.
- [18] Mohinder S. GREWAL et al. *Global positioning systems, inertial navigation, and integration*. Wiley-Interscience, jan. 2007, p. 553. ISBN : 9780470041901.

- [19] Jérôme Verdun HENRI DUQUENNE. "Le champ de pesanteur : notions fondamentales et méthodes modernes de détermination". In : *Cours aux étudiants PPMD*. ENSG, 2006.
- [20] Bernhard HOFMANN-WELLENHOF, Herbert LICHTENEGGER et Elmar WASLE. *GNSS Global Navigation Satellite Systems : GPS, GLONASS, Galileo, and more*. 1^{re} éd. Springer, déc. 2007. ISBN : 3211730125.
- [21] Bernhard HOFMANN-WELLENHOF et Helmut MORITZ. *Physical Geodesy*. 1^{re} éd. Springer, sept. 2005. ISBN : 3211235841.
- [22] J. IDHE, J. MAKKINEN et M. SACHER. "Conventions for the Definition and Realisation of a European Vertical Reference System (EVRS) - EVRS conventions 2007". In : 2008.
- [23] Len JACOBSON. *GNSS markets and applications*. Artech House, juin 2007, p. 240. ISBN : 9781596930421.
- [24] KLOBUCHAR. "A first order worldwide ionospheric time-delay algorithm". In : *Ionospheric Physics Laboratory, Air Force Cambridge Research Laboratories* (1975). Hanscom AFB, Massachusetts Klobuchar John A.
- [25] François L'ECU. "Calcul du quasi-geoïde QGF16 et de la grille de conversion altimétrique RAF16, état d'avancement et perspectives". In : XYZ (2017).
- [26] I.M. LONGMAN. "Formulas for Computing the Tidal Accelerations Due to the Moon and the Sun". In : *Journal of Geophysics Research* 64 (1959), p. 2351-2355.
- [27] J. Hinderer M. AMALVICT N. Debeglia. "The Absolute Gravity Measurements Performed By Sakuma In France, Revisited 20 Years Later". In : *Gravity and Geoid 3rd meeting of the IGGC* (2002), p. 76-83.
- [28] Ahmed EL-RABBANY. *Introduction to GPS : the Global Positioning System*. Artech House, 2002, p. 202. ISBN : 9781580531832.
- [29] M. SACHER et al. "EVRF2007 as Realization of the European Vertical Reference System". In : *Symposium oh the IAG sub-commision for Europe (EUREF) in Brussels, June 18-21 2008*. 2008.
- [30] Nel SAMAMA. *Global positioning : technologies and performance*. Wiley-Interscience, 2008, p. 440. ISBN : 9780471793762.
- [31] Wilfred SCHOFIELD et Mark BREACH. *Engineering surveying*. Butterworth-Heinemann, avr. 2007, p. 637. ISBN : 9780750669498.
- [32] SCINTREX. *User's guide : CG-3/3M Gravity Meter*. Scintrex Ltd. 1995.
- [33] Gunter SEEGER. *Satellite Geodesy*. 2 Revised. Walter de Gruyter, sept. 2003. ISBN : 3110175495.
- [34] Victor G. SZEBEHELY et Hans MARK. *Adventures in celestial mechanics*. Wiley-VCH, 1998, p. 328. ISBN : 9780471133179.
- [35] Wolfgang TORGE. *Geodesy*. 3rd ed., De Gruyter, 2001, p. 416.
- [36] Wolfgang TORGE. *Gravimetry*. Walter de Gruyter, déc. 1989, p. 488. ISBN : 9783110107029.
- [37] James Bao-yen TSUI. *Fundamentals of global positioning system receivers : a software approach*. John Wiley et Sons, 2005, p. 373. ISBN : 9780471706472.
- [38] P. WESSEL et W.H.F. SMITH. "Free software helps map and display data". In : *EOS Trans. Amer. Geophys. U.* 72 (1991), p. 445-446.
- [39] P. WESSEL et W.H.F. SMITH. *The Generic Mapping Tools Technical Reference and Cookbook*. SOEST-NOAA. 2010.
- [40] Derek Van WESTRUM. *A10 absolute portable gravimeter*. Micro-g Solutions, Inc., 2005.
- [41] Guochang XU. *GPS : theory, algorithms, and applications*. Springer, 2007, p. 354. ISBN : 9783540727149.

- [42] Guochang XU. *Orbits*. Springer, 2008, p. 236. ISBN : 9783540785217.

Table des figures

1.1	Exemple de clients Galigeo	5
1.2	Bureaux de Galigeo, 87 avenue d'Italie, 75013 Paris	6
1.3	Bureaux de Galigeo, 87 avenue d'Italie, 75013 Paris	7

Liste des tableaux

Annexes

A	Filtre de Kalman	23
B	Moindres carrés	25

FILTRE DE KALMAN

ANNEXE
A

Contenu de l'annexe sur Kalman...

MOINDRES CARRÉS

ANNEXE
B

Contenu de l'annexe sur MC...