



MASTÈRE SPÉCIALISÉ® SCIENCE DES DONNÉES POUR LA CONNAISSANCE CLIENT

Année Scolaire 2025/2026

Table des matières

UE 1 : HARMONISATION	5
RAPPEL DE MATHÉMATIQUES ET DE STATISTIQUE	5
PRINCIPES DE MODÉLISATION, STATISTIQUES MULTIVARIÉES	6
ÉCONOMÉTRIE	7
SÉRIES TEMPORELLES	8
PYTHON ORIENTÉ OBJET	10
UE 2 MACHINE LEARNING	11
APPRENTISSAGE STATISTIQUE	11
APPRENTISSAGE PROFOND	12
WEBMINING ET NATURAL LANGUAGE PROCESSING	13
APPRENTISSAGE PAR RENFORCEMENT	15
TRAITEMENT DU LANGAGE AVANCÉ	16
OPTIMISATION POUR L'APPRENTISSAGE	17
UE 3 MODELISATION	18
ÉCONOMÉTRIE DES DONNÉES DE PANEL	18
MODÈLES DE RÉGRESSION BAYÉSIENNE	20
STATISTIQUE ET ÉCONOMÉTRIE SPATIALE	21
MODÈLES DE CHOIX DISCRETS	22
UE 4 PROFESSIONNALISATION 1	23
STRATEGIE MARKETING AVEC DRYRUN MARKETING	23
INTRO ET PRÉSENTATION DU MARKETING & CRM	24
MARKETING EXPÉRIENTIEL	25
MARKETING DIGITAL	26
DROIT ET ÉTHIQUE DES DONNÉES	27
UE 5 PROFESSIONNALISATION 2	28
MESURE ET ANALYSE DE LA SATISFACTION	28
SCORING	29
STATISTIQUE EXPLORATOIRE, DATA VISUALISATION ET DATA STORY TELLING	30
CLOUD	31

Présentation du Mastère

INTITULÉ UNITÉ / MODULE	INTITULÉ COURS	VOLUME HORAIRES EVOLUTION	Langue d'enseignement	CRÉDITS ECTS AFFECTÉS
UE 1 "UE Harmonisation ENSAI"		117		10
MS1-01	Rappel de mathématiques et de statistique	36	Français	3
MS1-02	Principes de modélisation, statistiques multivariées	30	Français	3
MS1-03	Économétrie	15	Français	2
MS1-04	Séries temporelles	24	Français	1
MS1-05	Python orienté objet	12	Français	1
UE 2 "UE Machine Learning"		108		12
MS2-01	Apprentissage statistique	30	Français	4
MS2-02	Apprentissage profond	24	Français	2
MS2-03	Webmining et natural language processing	21	Français	2
MS2-04	Apprentissage par renforcement	9	Français	1
MS2-05	Traitements du langage avancé	12	Français	1
MS4-06	Optimisation pour l'apprentissage	12	Français	2
UE 3 "UE Modélisation"		42		5,5
MS3-01	Économétrie des données de panel	12	Français	1,5
MS3-02	Modèle de Régression bayésienne	12	Français	1,5
MS3-03	Statistiques et économétrie spatiale	12	Français	1,5
MS3-04	Modèles de choix discrets	6	Français	1
UE 4 "UE de professionnalisation 1"		59		8
MS4-01	Introduction, présentation du marketing et CRM	15	Français	2
MS4-02	Stratégie marketing avec Dryrun Marketing	23	Français	2
MS4-03	Marketing expérientiel	12	Français	2
MS4-04	Marketing digital	6	Français	1
MS4-05	Droit et éthique des données	3	Français	1
UE 5 "UE de professionnalisation 2"		77		9,5
MS4-01	Scoring	24	Français	3
MS4-02	Mesure et analyse de la satisfaction (appli text mining, ...)	17	Français	2
MS4-03	Cloud	21	Français	2
MS4-04	Statistique exploratoire, Data visualisation et data story telling	15	Français	2,5
TOTAL		403		45
Séminaire	IA et prise de décision	18	Français	0
UE 6 "UE Mission entreprise"				10
Mission "entreprise"			Français	10
UE 7 "UE Thèse/mémoire"				20
Thèse/mémoire			Français	20
TOTAL				30

UE 1 : HARMONISATION

RAPPEL DE MATHÉMATIQUES ET DE STATISTIQUE

Enseignant : Claude PETIT
ECTS : 3
Volume horaire : 36 heures (Cours magistral + TD)

Objectif pédagogique

- Revoir (ou voir) les notions de statistiques mathématiques et les outils probabilistes indispensables pour suivre sereinement les cours du Mastère Spécialisé.
- Pratiquer sur des exercices.
- Avoir un aperçu succinct de certains cours à venir.

Contenu de la matière

- Modèles statistiques et estimation ponctuelle (différents estimateurs, biais, erreur quadratique, propriétés asymptotiques, maximum de vraisemblance, cadre bayésien).
- Estimation optimale (exhaustivité, statistique libre et complète, information de Fisher, optimalité).
- Intervalles de confiance (intervalle et région de confiance, fonction pivotale, différentes méthodes pour construire un intervalle de confiance, intervalle de crédibilité dans le cadre bayésien).
- Tests statistiques (tests paramétriques, risque, niveau, puissance, efficacité, quelques tests non paramétriques).

Pré-requis

- Les cours de probabilités et de théorie de la mesure de L3 ou de première année d'école d'ingénieurs.

Contrôle des connaissances

- Devoir sur table + projet + rendu d'exercices

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

PRINCIPES DE MODÉLISATION, STATISTIQUES MULTIVARIÉES

Enseignants : Brigitte GELEIN (ENSAI)

ECTS : 3

Volume horaire : 30 heures (Cours magistral + TD)

Objectif pédagogique

A l'issue de cet enseignement, les élèves devront être capables de procéder à un choix de modèle à partir d'un jeu de données et d'une problématique, de mettre en œuvre les méthodes d'analyse exploratoire multivariée ainsi que certaines méthodes de classification, d'interpréter et de présenter les résultats en vue d'un échange avec des non-statisticiens

Contenu de la matière

- Principe de construction de modèle (régression linéaire et logistique)
- Validation du modèle et capacité de prédiction
- Analyses factorielles
- Méthodes de classification (géométriques et probabilistes)

Pré-requis

- Non indiqué

Contrôle des connaissances

- Devoir sur table

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

ÉCONOMÉTRIE

Enseignant : Eric DELATTRE
ECTS : 2
Volume horaire : 15h (Cours magistral + TD)

Objectif pédagogique

Les données clients d'une entreprise ne sont pas le fruit d'expériences aléatoires, mais sont observées à partir de comportements humains (achats, clics sur des pages produits, réactions à des promotions...). Juger l'effet d'offres promotionnelles peut ainsi être difficile car se confondre avec des effets d'aubaine ou se limiter à une population spécifique (les jeunes très connectés par exemple). Dans un modèle statistique, cela se traduit par une corrélation entre les variables explicatives et le terme d'erreur (ce qu'on nomme endogénéité). L'objet de l'économétrie est d'évaluer un effet causal, en corrigeant cette endogénéité. Ce cours présente d'une part la démarche économétrique (spécification, identification, interprétation, lien avec la théorie économique et marketing), et d'autre part les modèles de base et les techniques d'inférence utilisés par les économètres (variables instrumentales notamment).

Contenu de la matière

À l'issue du cours, l'étudiant saura :

- Mobiliser les concepts de la théorie micro-économique et les techniques statistiques appropriées pour la construction d'indicateurs ou la conduite d'une étude économique ;
- Interpréter les résultats d'une étude économique, en distinguant approche descriptive et causale ;
- Analyser les sources et les effets de l'endogénéité des variables explicatives, et mettre en œuvre une analyse par variable instrumentale ;
- Identifier les biais et gains potentiels liés à l'usage de variables particulières : évaluation des politiques, données d'enquête, données de panel, données temporelles... ;
- Lire et discuter un texte économétrique avec un regard critique sur la qualité des données, le choix des méthodes et l'interprétation des résultats.

Pré-requis

- Non indiqué

Contrôle des connaissances

- Présentation

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

SÉRIES TEMPORELLES

Enseignant : Youssef ESSTAFA (Université Le Mans)
ECTS : 1
Volume horaire : 24 heures (Cours magistral + TD)

Objectif pédagogique

L'objectif principal du cours est la présentation des techniques d'analyse des séries temporelles multivariées couramment utilisées dans les applications. Deux cas sont distingués : séries stationnaires et séries non-stationnaires. Dans le cas stationnaire, il s'agit principalement de l'étude des modèles VAR (vectoriels auto-régressifs). Dans le cas non-stationnaire le cours se concentre sur l'inférence en présence de racines unités et à l'estimation de relations de cointégration. Le cours commence par un rappel de l'approche ARMA et des modèles (G)ARCH pour les séries univariées.

L'objectif de l'atelier est de mettre en œuvre les méthodes d'estimation et de test présentées dans la partie cours qui se déroule en parallèle. Les concepts théoriques seront illustrés par des données réelles traitées à l'aide du logiciel SAS.

Contenu de la matière

1. Processus stationnaires (rappels/extensions):
 - 1.1. Modélisation de l'espérance conditionnelle – les processus ARMA et extensions SARIMA.
 - 1.2. Modélisation de la variance conditionnelle – les processus (G)ARCH et leurs Extensions.
2. Processus non stationnaires univariés:
 - 2.1. Différentes formes de non stationnarité : tendance déterministe et racine unité.
 - 2.2. Tests de racine unité.
3. Processus stationnaires : les modèles VAR
 - 3.1. Stationnarité.
 - 3.2. Estimation et tests.
 - 3.3. Causalité et test de non causalité.
4. Processus non stationnaires: les processus à racine unité et cointégration.
 - 4.1. Cointégration – théorème de Granger et modèle à correction d'erreurs.
 - 4.2. Cointegration – modèle vectoriel à correction d'erreurs.

Contrôle des connaissances

Projet

Références bibliographiques

- BANERJEE, A., DOLADO, J.J., GALBRAITH, J.W., & HENDRY, D.F., 1993, *Co-integration, Error Correction, and the Econometric Analysis of non-stationary data*, Oxford University Press, New-York.
- BROCKWELL, P.J., & DAVIS, R.A., 1998, *Time Series : Theory and Methods*, Springer Series in Statistics, Springer Verlag, Second Edition.
- BROCKWELL, P.J., & DAVIS, R.A., 2002, *Introduction to Time Series and Forecasting (2nd ed.)*, Springer-Verlag, New-York.

- BOURBONNAIS,R., 2011, *Econométrie*, 8ème ed, Dunod.
- FRANCO, C.,& ZAKOIAN, J., 2009, *Modèles GARCH*, Economica.
- GOURIEROUX, G., 1992, *Modèles ARCH et applications financières*, Economica.
- GOURIEROUX, G., & MONFORT, A., 1995, *Séries temporelles et modèles dynamiques*, 2^{ème} ed., Economica.
- HAMILTON, J.D., 1994, *Time Series Analysis*, Princeton Univ. Press.
- MADDALA, G.S., & KIM, I.M., 1998, *Unit Roots, Cointegration and Structural Change*, Cambridge University Press.
- MIGNON, V., LARDIC, S., 2002, *Econométrie des séries temporelles macroéconomiques et financières*, Economica.
- PFAFF, B., 2008, *Analysis of Integrated and Cointegrated Time Series with R*, 2nd ed., Springer.
- TSAY, R.S, 2010, *Analysis of Financial Time Series*, 3rd edition, Wiley.

Langue d'enseignement

Français.

PYTHON ORIENTÉ OBJET

Enseignant : Gaëlle ROBARDET-NICOLAS
ECTS : 1
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

Remettre à niveau les étudiants sur les concepts et langages informatiques qui seront utilisés dans les UE du mastère : bases de données relationnelles (en SQL), algorithmique et programmation orientée objet (Python et UML).. Introduire le langage de R en vue de l'analyse et l'exploration des données et de la programmation numérique.

Contenu de la matière

- Effectuer des requêtes sur une base de données (en SQL)
- Concevoir une base de données adaptées à une application donnée.
- Programmer en Python.
- Concevoir (à l'aide d'UML) et implémenter (en Python) un logiciel composé de plusieurs classes d'objets réparties en divers bibliothèques.
- Les bases de R : l'environnement R, les objets, les fonctions de base
- Les données, les graphiques avec R
- Eléments de programmation

Documents pédagogiques

Envoyés par courriel avant la rentrée (pour bases de données).

Prérequis

Aucun

Contrôle des connaissances

TP noté

Références bibliographiques

- UML : <http://openclassrooms.com/courses/debutez-l-analyse-logicielle-avec-uml>
- Python : <https://openclassrooms.com/courses/apprenez-a-programmer-en-python>

Langue d'enseignement

Français

UE 2 MACHINE LEARNING

APPRENTISSAGE STATISTIQUE

Enseignant : Claude PETIT
ECTS : 4
Volume horaire : 30 heures (Cours magistral + TD)

Objectif pédagogique :

Ce cours propose une introduction aux grands principes de l'apprentissage supervisé et non-supervisé. L'objectif est de maîtriser les bases théoriques et de savoir utiliser les algorithmes des techniques classiques de l'apprentissage supervisé et quelques techniques de l'apprentissage non supervisé. Il doit permettre de connaître les cas d'usage des modèles, de les comparer, de les évaluer empiriquement et de savoir les implémenter en langage Python.

Contenu de la matière

Problématique des statistiques en grande dimension, malédiction de la dimension, dilemme biais/variance, théorie de la décision : oracle, classification binaire, régression, minimisation du risque, sur- et sous-apprentissage. Méthodes de partitionnement : algorithme des plus proches voisins, arbres de décision et algorithme CART. Méthodes de convexification : machines à vecteurs de support (SVM), méthodes à noyaux. Méthodes d'agrégation: Boosting, Bagging, Stacking, forêts aléatoires. Courte introduction à l'apprentissage non supervisé: clustering et estimation de densités. Apprentissage semi-supervisé : cas des GNN (Graph Neural Networks). Les techniques de bootstrap et l'algorithme de descente de gradient seront présentés en travaux pratiques.

Pré-requis

un cours de statistique mathématique et des notions de programmation en Python. Des connaissances de base en algèbre linéaire et optimisation

Contrôle des connaissances

projet informatique.

Références bibliographiques

- Introduction to Statistical Learning with Application in Python, de James, Witten, Hastie, Tibshirani, Taylor, aux éd. Springer.
- Introduction à l'apprentissage automatique, de Frédéric Sur, polycopié de l'école des Mines de Nancy.
- Mathematics for Machine Learning, de G. Thomzd, polycopié de l'université de Berkeley.
- The Elements of Statistical Learning, de Hastie, Tibshirani, Friedman, aux éd. Springer.
- The Nature of Statistical Learning, Vladimir Vapnik, aux éd. Wiley.

Langue d'enseignement

Français

APPRENTISSAGE PROFOND

Enseignant : Johan FAOUZI (ENSAI)
ECTS : 2
Volume horaire : 24 heures (Cours magistral + TD)

Objectif pédagogique

De nombreuses tâches d'apprentissage et de traitement du signal visent à retrouver un ensemble de grandeurs inconnues (état d'un système, modèle génératif, etc.) à partir de données. Malheureusement, dans de nombreuses situations, les données disponibles s'avèrent insuffisantes pour lever l'ambiguité sur les quantités à inférer ou les estimer avec une précision suffisante. Une manière classique de contourner ce problème consiste à introduire une information « *a priori* » sur la solution recherchée. Plus particulièrement, dans ce cours nous montrons comment lever l'ambiguité inhérente à certains problèmes en « pénalisant » les solutions ne présentant pas certaines caractéristiques d'intérêt. L'objectif de ce cours est d'identifier et manipuler les méthodes de pénalisation les plus courantes.

Contenu de la matière

- Identifier la pénalisation la plus adaptée à une tâche.
- Résoudre un problème d'optimisation comportant un terme de régularisation.
- Régler les paramètres du modèle

Prérequis

- Algèbre linéaire
- Optimisation convexe
- Programmation en Python

Contrôle des connaissances

TP notés + examen final

Références bibliographiques

- C. Bishop. *Pattern recognition and machine learning*. Springer-Verlag New York, 2006.
S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing. Applied and Numerical Harmonic Analysis*. Birkhäuser, 2013.
D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, USA, 2003.

Langue d'enseignement

Français

WEBMINING ET NATURAL LANGUAGE PROCESSING

Enseignants : Cyrielle MALART (IRISA) + Christian RAYMOND
ECTS : 2
Volume horaire : 21 heures (Cours magistral + TD)

Objectif pédagogique

- Collecter, nettoyer et structurer des données textuelles
- Comprendre les fondements linguistiques et statistiques des approches modernes en TAL
- Comprendre le paysage actuel pour choisir une méthode de traitement automatique de la langue pour une tâche classique (classification, analyse de sentiment, détection d'entités, etc.)
- Connaitre les outils logiciels standards du domaine (spaCy, Stanza, nltk, etc.) et savoir choisir pour son cas d'utilisation

Contenu de la matière

1. What is natural language and why process it
2. Representation of words: symbolic vs. distributional approaches
3. Linguistic aspects: tagging and parsing
4. Representation and classification of documents
5. Language modeling and contextual word representation
6. Sequence-to-sequence models and attention
7. Standard NLP tasks today

Prérequis

- Bases de programmation Python (boucles, structures, fonctions)
- Statistiques et probabilités de base (distributions, estimation, Bayes)
- Notions élémentaires en classification supervisée (logistique, SVM, cross-validation)
- aisance avec la lecture en anglais (documentation, articles, corpus)

Contrôle des connaissances

Projet en binôme (rapport) au cours des ateliers 3 et 4

Références bibliographiques

- Daniel Jurafsky, James H. Martin. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition, 2nd edition, Prentice-Hall, 2009.
 - Draft of the 3rd edition partly available at <https://web.stanford.edu/~jurafsky/slp3>.
- Yoav Goldberg. Neural Network Methods for Natural Language Processing. 2017.
 - An earlier draft is freely available online at <http://u.cs.biu.ac.il/~yogo/nnlp.pdf>.
- Kevin Gimpel's lectures (Toyota Technological Institute at Chicago and UChicago) on :
 - Natural Language Processing (<https://ttic.uchicago.edu/~kgimpel/teaching/31190-s18/index.html>) and
 - Advanced Natural Language Processing (<https://ttic.uchicago.edu/~kgimpel/teaching/31210-s19/index.html>).
- Chris Manning's CS224n Lecture Notes and Videos (Stanford)
 - Website for the course (scroll to the bottom for the slides and lecture notes) : <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1194/>

- Youtube playlist of the lectures :
<https://www.youtube.com/playlist?list=PLoROMvodv4rOhcuXMZkNm7j3fVwBBY42z>
- Jay Alammar's 2018 blog post about transformers :
<https://jalammar.github.io/illustrated-transformer/>

Langue d'enseignement Anglais

APPRENTISSAGE PAR RENFORCEMENT

Enseignants : Mohammadreza MOUSAVI KALAN (ENSAI)

ECTS : 1

Volume horaire : 9 heures (Cours magistral + TD)

Objectif pédagogique

Objectif 1 : savoir modéliser une situation comme un problème d'apprentissage par renforcement (choix d'une fonction de récompense, d'un espace d'état, d'un espace d'actions...)

Objectif 2 : être capable d'apprendre une politique à partir de données et de l'évaluer

Objectif 3 : comprendre les limites de la modélisation par apprentissage par renforcement

Contenu de la matière

- Modèles de bandits
- Dilemme exploration-exploitation
- Algorithmes de bandits
- Regret
- Processus de décision markovien (MDP)
- Équation de Bellman
- Fonction état-valeur
- Algorithmes pour MDP

Prérequis

Apprentissage profond et Chaînes de Markov facilitent la compréhension.

Contrôle des connaissances

Devoir maison

Références bibliographiques

- Tor Lattimore and Csaba Szepesvári, Bandit Algorithms, 2020 (<https://tor-lattimore.com/downloads/book/book.pdf>)
- Sutton and Barto, Introduction to Reinforcement Learning, 2020 (<http://incompleteideas.net/book/RLbook2020.pdf>)
- David Silver course on RL: <https://www.davidsilver.uk/teaching/>
- GYM, OpenAI, <https://gym.openai.com/>

Langue d'enseignement

Français.

TRAITEMENT DU LANGAGE AVANCÉ

Enseignants : Cyrielle MALART (IRISA)
ECTS : 1
Volume horaire : 12 heures (TD)

Objectif pédagogique

Entraîner un nouveau modèle de langue

Contenu de la matière

1. Document preprocessing
2. Applied neural networks
3. Long- and short-term memory models (LSTM)
4. How to apply an embedding
5. Entity recognition as token-sequence labelling

Pré-requis

Programmation avec Python, Apprentissage statistique, Traitement automatique de la langue

Contrôle des connaissances

TP noté

Références bibliographiques

- Daniel Jurafsky, James H. Martin. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition, 2nd edition, Prentice-Hall, 2009. Draft of the 3rd edition partly available at <https://web.stanford.edu/~jurafsky/slp3>.
- Yoav Goldberg. Neural Network Methods for Natural Language Processing. 2017. An earlier draft is freely available online at <http://u.cs.biu.ac.il/~yogo/nlp.pdf>.
- Kevin Gimpel's lectures (Toyota Technological Institute at Chicago and UChicago) on Natural Language Processing (<https://ttic.uchicago.edu/~kgimpel/teaching/31190-s18/index.html>) and on Advanced Natural Language Processing (<https://ttic.uchicago.edu/~kgimpel/teaching/31210-s19/index.html>).

Langue d'enseignement

Français.

OPTIMISATION POUR L'APPRENTISSAGE

Enseignant : Mohammadreza Mousavi Kalan (Ensaï)
ECTS : 2
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

Les systèmes de recommandation choisissent les options à présenter à des utilisateurs parmi un grand nombre de possibilités. Ils permettent par exemple de recommander la prochaine vidéo à regarder, le prochain morceau à écouter, les photos à montrer... Le cours présentera les modèles utilisés pour construire de tels systèmes.

Contenu de la matière

- Objectif des systèmes de recommandation
- État de l'art des systèmes de recommandation (plus proches voisins, filtrage collaboratif...)
- Évaluation des systèmes de recommandation
- Problèmes rencontrés par les systèmes de recommandation et solutions afférentes (démarrage à froid, compromis exploration-exploitation...)

Compétences

- Mise en application des modèles étudiés
- Évaluation d'un système de recommandation

Pré-requis

Python, principes de l'apprentissage automatique

Contrôle des connaissances

Devoir sur table

Références bibliographiques

- Statistical Methods for Recommender Systems. Deepak K. Agarwal, Bee-Chung Chen. 2016.
- Recommender Systems: The Textbook. Charu C. Aggarwal. Springer, 2016.
- Bandit algorithms for Website optimization. John Myles White. O'Reilly Media.
- Blog et tutoriels de Sébastien Bubeck's : <https://blogs.princeton.edu/imabandit/>
- Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. Sébastien Bubeck, Nicolò Cesa-Bianchi. <https://arxiv.org/abs/1204.5721>
- Bandit Algorithms. Tor Lattimore and Csaba Szepesvári. Cambridge University Press. <https://tor-lattimore.com/downloads/book/book.pdf>

Langue d'enseignement

Français

UE 3 MODELISATION

ÉCONOMÉTRIE DES DONNÉES DE PANEL

Enseignant : Bertille PICARD (ENSAI)
ECTS : 1,5
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

L'objectif de ce cours est de fournir les outils utiles à l'estimation de modèles économétriques sur données de panel. Ce cours décrit les méthodes et leur mise en application pratique. Des exemples numériques ainsi que des travaux de recherche empiriques sont présentés tout au long du cours pour illustrer l'utilisation de ces méthodes.

- * Introduction
 - * Illustration
 - * Définitions
 - * Notations
 - * Ecritures du modèle de panel
 - * Les opérateurs
- Modèle linéaire statique
 - * Le modèle à effets fixes
 - * Le modèle à effets aléatoires
- * Le modèle linéaire à effets individuels corrélés
 - * L'approche de Mundlak et Chamberlain
 - * Les méthodes de variables instrumentales
 - * L'estimation du modèle dynamique de panel
 - * Exemples
- * Extensions
 - * Le modèle à coefficients aléatoires
 - * Modèles de panel non linéaires
- * Analyses d'articles
 - * Les étudiants devront lire un article d'économétrie appliquée qui reprendra et prolongera les différentes méthodes du cours. Les thèmes abordés seront l'économie de la santé, l'économie spatiale et le marketing

Pré-requis

Micro-économétrie avancée.

Contrôle des connaissances

Examen écrit.

Références bibliographiques

- * B. H. BALTAGI. *Econometric Analysis of Panel Data* (2nd ed.), J. Wiley, 2001

Campus de Ker Lann, 51 rue Blaise Pascal - BP37203 - 35172 BRUZ CEDEX
Tél : 33 (0)2 99 05 32 32 / scolarite@ensai.fr
www.ensai.fr

- * P. SEVESTRE. *Économétrie des données de panel*, Dunod, 2002
- * J. M. WOOLDRIDGE. *Econometric Analysis of Cross Section and Panel Data (2nd ed.)*, 2010, MIT Press
- * A. PIROTTÉ. *Économétrie des données de panel*, Economica, 2011

Langue d'enseignement

Français

MODÈLES DE RÉGRESSION BAYÉSIENNE

Enseignant : Denys POMMERET (AMU – Aix-Marseille Université)
ECTS : 1,5
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

Le cours présentera la théorie de l'inférence Bayésienne. Des exemples et exercices seront développées pendant le cours.

A l'issue de ce cours, les étudiants doivent être capable de :

- calculer la distribution à-posteriori pour des modèles conjugués, construire un estimateur Bayésien, construire test d'hypothèses Bayésiens et régions de confiance ;
- comprendre la différence entre une procédure d'inférence fréquentiste et Bayésienne
- implémenter sur ordinateur une procédure d'inférence Bayésienne ;
- résoudre des exercices portants sur les notions apprises pendant le cours

Contenu de la matière

- Les principes Bayésiens, échangeabilité, principe de vraisemblance.
- Détermination de lois à priori
- Inférence Bayésienne : estimation ponctuelle, régions de confiance, tests d'hypothèses.
- Propriétés asymptotiques des approches Bayésiennes.
- Choix et comparaison de modèles.

La mise en application de ces méthodes est faite en atelier sous R.

Pré-requis

Cours de statistique inférentielle et statistique computationnelle.

Contrôle des connaissances

Devoir maison

Références bibliographiques

Non indiqué

Langue d'enseignement

Français

STATISTIQUE ET ÉCONOMÉTRIE SPATIALE

Enseignant : Salima BOUAYAD AGHA (Université Le Mans)
ECTS : 1,5
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

Après avoir passé en revue les différents types de données spatiales, le cours présente les outils de base de la statistique spatiale qui permettent de mesurer le degré de signification statistique des configurations et des relations spatiales de données géoréférencées, qui vont ainsi compléter et enrichir l'approche strictement cartographique.

Le cours s'attache ensuite plus spécifiquement à l'étude des données économiques. Les méthodes de l'économétrie spatiale sont de plus en plus utilisées dans de nombreux domaines (croissance, économie régionale et urbaine, marketing, étude des marchés immobiliers, ...). Favorisées par le développement des systèmes d'information géographique qui permettent de disposer simultanément des valeurs prises par les variables d'intérêt et de leur localisation géographique, ces méthodes permettent de prendre en compte dans la modélisation les phénomènes d'interaction spatiale de différentes manières.

Après avoir présenté les différentes manières de formaliser les effets spatiaux (effet de débordement et de dépendance spatiale, hétérogénéité) seront exposés les différentes spécifications économétriques spatiales ainsi que leur estimation par différentes méthodes (maximum de vraisemblance et méthode des moments généralisés). Les tests de spécifications les plus courants seront également exposés. Les exposés seront illustrés par des exemples issus de la littérature récente dans ce domaine.

De nombreux exemples utilisant R ou STATA illustrent les sujets abordés. Le cours sera complété de 3 ateliers, l'un sur la cartographie et les méthodes exploratoires des données, les 2 autres sur l'économétrie.

Contrôle des connaissances

Devoir sur table + devoir maison

MODÈLES DE CHOIX DISCRETS

Enseignant : Marion GOUSSE (ENSAI)
ECTS : 1
Volume horaire : 6 heures (Cours magistral + TD)

Objectif pédagogique

Ce cours a pour objectif principal de présenter les avancées récentes en matière de modélisation des choix discrets.

Après quelques rappels quant à la structure des modèles économétriques de choix discret (modèles à utilités aléatoires) et au modèle de référence employé par les économètres (le modèle Logit Multinomial), le cours vise à tout d'abord à présenter deux extensions essentielles des modèles de choix discret de référence : (a) les modèles de choix discret à paramètres aléatoires et (b) les modèles de choix discret avec effets d'agglomération et/ou de congestion. Ces modèles sont employés dans tous les domaines utilisant la microéconométrie appliquée : choix des modes et des voies de transport, choix de localisation d'activité ou d'habitation, modélisation des parts de marchés en situation de concurrence imparfaite, choix de sites récréatifs, ...

Le cours abordera les questions liées à la spécification et à l'estimation de ces modèles, en théorie et à partir d'exemples traités en détail.

Contenu de la matière

1. Rappels
 - a. Modèles à utilités aléatoires
 - b. Le modèle Logit Multinomial
 - c. Cas des grands ensembles de choix
2. Modèles à paramètres aléatoires
 - a. Intérêt et spécification des modèles à paramètres aléatoires
 - b. Loi mélangeante continue : estimation par le MV
 - c. Loi mélangeante discrète : estimation par le MV via l'algorithme EM
 - d. Quelques aspects pratiques
3. Modèles de localisation avec effets de congestion et/ou d'agglomération
 - a. Effets de congestion et d'agglomération et équilibres
 - b. Spécification et estimation des modèles de choix discrets avec effets de congestion ou d'agglomération
4. Supplément. Modèles de parts de marché

Pré-requis

Microéconométrie appliquée, Micro-économie, Microéconométrie avancée.

Contrôle des connaissances

Examen écrit.

UE 4 PROFESSIONNALISATION 1

STRATEGIE MARKETING AVEC DRYRUN MARKETING

Enseignant : Rémi TRIOLET (Simulation Factory)

ECTS : 2

Volume horaire : 23 heures (Cours magistral)

Objectif pédagogique

Ce cours a pour objectif de présenter les bases du marketing et son évolution depuis 50 ans. L'objectif est de donner un cadre conceptuel aux data scientist leur permettant de comprendre les enjeux du traitement des données et d'utiliser au mieux leurs compétences.

Contenu de la matière

- Comprendre le marketing
- Analyser le marché
- Elaborer une stratégie marketing
- Construire l'offre de marché
- Gérer des plans d'action marketing

Contrôle des connaissances

Etude de cas

Références bibliographiques

- P. Kotler, K. Keller, D. Manceau, B. Dubois, *Marketing management*, Pearson Education, 2009.

Langue d'enseignement

Français.

INTRO ET PRÉSENTATION DU MARKETING & CRM

Enseignant : Catherine HEITMAN MENAGER (Université de Rennes)
ECTS : 2
Volume horaire : 15 heures (Cours magistral + TD)

Objectif pédagogique

Customer Relationship Management (CRM) increasingly emphasized since the 90's is considered by some academics as a paradigm shift in marketing as it puts the focus on the building of long term relationships, a notion that cannot be captured through the traditional marketing-mix concept. Students will gain insight into this critical area of marketing management through discussions readings, lectures and a small group project or case study.

3 MAJOR OBJECTIVES :

1. Provide a rigorous definition and thorough understanding of the key underlying concepts of CRM, for instance: relationships, satisfaction, loyalty, and commitment. Students will be able to distinguish CRM from Direct Marketing, a method with which it is often confused.
2. Understand how services and relationships cross-fertilize each other. Relationships materialize in a different manner between products and services. In the case of services, a major issue is the service relationship that exists even in a single transaction. This observation puts the focus on the quality the customer experience, the different channels of interaction and the role of contact staff, considered as a part-time marketer. Whatever the nature of the core offering – product or service – developing a relationship requires developing services.
3. Present different methods and tools key to implementing a relationship marketing strategy. For instance, the development of individual customer knowledge enabling one to assess Customer Lifetime Value (CLV), customer relationship lifecycle and other Relationship Marketing Instruments (RMI) such as loyalty programs.

Contenu de la matière

Il sera fourni par l'intervenante.

Pré-requis

Aucun.

Contrôle des connaissances

50% Participation & Discussion – Individual
50% Small Team Project / Case Study with oral presentation

Références bibliographiques

Seront fournies par l'enseignante

Langue d'enseignement

Français

MARKING EXPÉRIENTIEL

Enseignant : Bruno DAUCE (Université d'Angers)
ECTS : 2
Volume horaire : 12 heures (Cours magistral + TD)

Objectif pédagogique

Ce cours a pour objectif de présenter un panorama large des évolutions que connaît aujourd'hui la fonction marketing. Dans un premier temps, nous ferons un bref rappel sur la démarche marketing avant de présenter quelques éléments du comportement du consommateur. Nous nous attacherons également à essayer de comprendre comment évolue la consommation et à identifier les lignes directrices qui guident le comportement du consommateur. Enfin, nous présenterons les mutations que connaît aujourd'hui le marketing en insistant sur le développement du marketing sensoriel et sur l'impact d'Internet sur la stratégie marketing.

Contenu de la matière

1. Introduction : rappels sur la démarche marketing - Illustrations : petit bateau, Puma
2. Le consommateur et son comportement
3. Le consommateur d'hier, d'aujourd'hui et de demain
4. Du marketing expérientiel au marketing sensoriel
5. E-marketing

Pré-requis

Marketing mix.

Contrôle des connaissances

Devoir sur table.

Références bibliographiques

- J. LENDREVIE, J. LEVY, D.LINDON. *Mercator : Théorie et pratique du marketing* (9^e éd.), 2009
- P. KOTLER, B. DUBOIS, D. MANCEAU. *Marketing management* (14^e éd.), 2012
- J.-F. LYOTARD. *Le postmoderne expliqué aux enfants*, Galilée, 2005
- M. MAFFESOLI. *Le temps des tribus* (3^e éd.), Le livre de Poche, 2000
- R. ROCHEFORT. *La société des consommateurs*, Éditions Odile Jacob, 1995
- R. ROCHEFORT. *Le consommateur entrepreneur*, Éditions Odile Jacob, 1997
- S. RIEUNIER. *Marketing sensoriel du point de vente : Créer et gérer l'ambiance des lieux commerciaux* (2^e éd.), Dunod, 2006
- A. GIBOREAU et L. BODY. *Le marketing sensoriel : De la stratégie à la mise en œuvre*, Vuibert, 2007
- N. GUEGUEN. *100 petites expériences en psychologie du consommateur : Pour mieux comprendre comment on vous influence*, Dunod, 2005

Langue d'enseignement

Français.

MARKETING DIGITAL

Enseignant : David FLOURIOT (Avanci)
ECTS : 1
Volume horaire : 6 heures (Cours magistral + TD)

Objectif pédagogique

En plus de renforcer leurs connaissances marketing indispensables à la compréhension des contextes professionnels dans lesquels ils vont évoluer, l'objectif de ce séminaire est avant tout de sensibiliser les étudiants aux nouvelles approches marketing liées aux canaux dits « digitaux » (site web et e-commerce, email, réseaux sociaux, mobile) et les complexités techniques adjacentes. Cet enseignement devra leur permettre d'être en pleine connaissance des KPI (Indicateurs Clé de Performance) qui servent de base de travail aux analyses de connaissance clients et mesure de la performance business.

Contenu de la matière

Sera fourni par l'enseignant.

Pré-requis

Marketing mix.

Contrôle des connaissances

Participation

Références bibliographiques

Seront fournies ultérieurement.

Langue d'enseignement

Français.

DROIT ET ÉTHIQUE DES DONNÉES

Enseignant : Stéphane ASTIER
ECTS : 1
Volume horaire : 3 heures (Conférence)

Objectif de la matière

L'objectif est d'apporter aux étudiants une « culture générale » juridique leur permettant, lorsqu'ils seront en poste, d'échanger plus facilement avec leurs services juridiques sur les aspects spécifiques de la réglementation « informatique et libertés ».

En s'appuyant sur l'actualité (géopolitique, technologique, économique), ce séminaire mettra en perspective une série d'enjeux juridiques et éthiques liés à la digitalisation de l'ensemble de la société tant dans l'environnement professionnel que personnel.

Des exemples concrets de procédures en cours devant les instances françaises et internationales ainsi que des faits d'actualité viendront illustrer cette présentation et alimenter une réflexion globale sur l'impact de l'évolution technologique sur notre civilisation.

Pré-requis

- Non indiqué

Contrôle des connaissances

Devoir sur table

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

UE 5 PROFESSIONNALISATION 2

MESURE ET ANALYSE DE LA SATISFACTION

Enseignants : Anne-Cécile GAY

ECTS : 2

Volume horaire : 17 heures (Cours magistral + TD)

Objectif pédagogique

La satisfaction client est devenue un enjeu crucial pour toutes les entreprises et son suivi est primordial. Le cours permet de :

- Connaitre les différentes approches conceptuelles de la satisfaction et de l'insatisfaction du consommateur ou de l'usager.
- Comprendre et interpréter les différents indicateurs qui mesurent la satisfaction et l'insatisfaction.
- Mesurer le lien entre satisfaction/insatisfaction et comportement du consommateur.
- Utiliser les nouveaux outils du NLP (Natural Language processing) pour traiter les informations textuelles et d'image collectées en particulier sur les sites internet via les outils numériques.

Contenu de la matière

- Principales théories de la satisfaction
- Les indicateurs classiques de la mesure de la satisfaction
- Outils informatiques du NLP, analyse textuelle, analyse sémantique, analyse des sentiments

Pré-requis

- Non indiqué

Contrôle des connaissances

- étude de cas

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

SCORING

Enseignant : Jean-Philippe KIENNER (JPK Conseil et Formation)
ECTS : 3
Volume horaire : 24 heures (Cours magistral + TD + pro)

Objectif pédagogique

L'objectif de ce cours est de présenter aux élèves l'une des méthodologies phares dans les études statistiques appliquées au marketing : le scoring.

En plus de l'objectif « technique » de maîtrise de la méthodologie de scoring, l'accent sera mis sur les problèmes fréquemment rencontrés en entreprise : population éligible, période d'observation, latence, nettoyage des données, restitution aux non-statisticiens...

Le fil rouge du cours consistera en la mise en œuvre complète d'un score et sera réalisé sur une base de données clients.

Contenu de la matière

1. Panorama des études réalisées
2. Principes du scoring
3. Base d'étude
 - 3.1. Population éligible
 - 3.2. Variable à expliquer
 - 3.3. Date de référence
 - 3.4. Échantillons d'apprentissage et de validation
 - 3.5. Variables explicatives
4. Modélisation
 - 4.1. Techniques de modélisation
 - 4.2. Construction des modèles
 - 4.3. Comparaison des modèles
 - 4.4. Interprétation du modèle final
5. Utilisation du score

Pré-requis

Régression sur variable catégorielle, statistique descriptive, statistique inférentielle

Contrôle des connaissances

Projet

Références bibliographiques

- G. SAPORTA. *Probabilités, analyse des données et statistique* (3^e éd.), 2011
- S. TUFFÉRY. *Data Mining et statistique décisionnelle* (3^e éd.), 2010
- S. TUFFÉRY. *Étude de cas en statistique décisionnelle*, 2009

Langue d'enseignement

Français.

STATISTIQUE EXPLORATOIRE, DATA VISUALISATION ET DATA STORY TELLING

Enseignant : Anil PALALI (Unlimitail)
ECTS : 2,5
Volume horaire : 18 heures (Cours magistral + TD)

Objectif pédagogique

Faire découvrir aux étudiants comment raconter des histoires passionnantes à partir de données brutes.

Portée à l'origine par la multiplication des données présentes sur internet et leur exploitation, la dataviz s'est étendu hors du monde numérique. Elle réinterroge la manière de communiquer les résultats au sein d'une entreprise, de produire de l'information dans les médias mais aussi d'exploiter les résultats d'une recherche expérimentale. La dataviz se place au cœur de toutes les interfaces.

Ce cours aborde les méthodes utilisées pour construire un scénario solide de communication grâce aux différentes formes de visualisation de données. Il montrera comment placer l'expérience utilisateur au cœur de la réflexion d'une restitution efficace. Centrées sur la production des sujets, les notions abordées sont largement illustrées par des exemples et des contre-exemples.

Le cours sera comporte un tiers de séances magistrales et deux tiers de séances pratiques

Contenu de la matière

A l'issue du cours, l'étudiant saura

- placer des données dans un contexte politique, social et dans le contexte métier du marketing et en tirer parti pour raconter une histoire plus précise
- éviter les erreurs à ne pas commettre lorsque l'on représente des données, en termes de représentation schématique et visuelle de l'information
- fabriquer une dataviz simple à partir d'un jeu de données public, en Javascript
- identifier les animations et les représentations de données les plus adaptées au message qu'il souhaite faire passer

Pré-requis

- Non indiqué

Contrôle des connaissances

- projet

Références bibliographiques

- Non indiqué

Langue d'enseignement

- Français

CLOUD

Enseignant : Rémi PEPIN (INSEE)
ECTS : 2
Volume horaire : 21 heures (Cours magistral + TD)

Objectif pédagogique

- Différencier le *cloud computing* du *on-premises computing*
- Sélectionner le meilleur type de service entre IaaS, PaaS, SaaS pour un besoin particulier
- Provisionner des services divers dans le cloud
- Manipuler différents services de base dans le cloud (stockage, calcul, base de données)
- Déployer une architecture hautement disponible dans le cloud

Contenu de la matière

Depuis l'arrivée en 2006 de l'offre commerciale d'*Amazon Web Service* (AWS) comportant un service de stockage et de calcul sur le marché, le monde de l'informatique s'est drastiquement métamorphosé. Plus besoin pour une entreprise d'investir dans un *datacenter*, il lui suffit en quelques clics de louer des machines pour répondre à ses besoins. Quinze ans après, AWS propose désormais plus de 200 services gérés, et n'est plus seul sur le marché (citons *Gougle Cloud Plateform* (GCP), *Microsoft Azur*, OVH).

S'il est plus nécessaire d'investir dans des machines physiques, il faut néanmoins comprendre comment fonctionnent les services proposés par un *Cloud Provider* et comment les faire communiquer entre eux. Le cours va ainsi aborder les notions suivantes :

- Le stockage objet
- Les instances de calcul
- Les bases de données
- Les règles basiques de sécurités dans le cloud
- La création de réseau
- Le *serverless*
- L'*Infrastructure as Code*

À la fin du cours, les élèves seront capables de déployer une architecture hautement disponible dans le cloud.

Pré-requis

Python niveau initiation, BdD niveau initiation

Contrôle des connaissances

Quizz, mini projet

Références bibliographiques

AWS Cookbook. (2022). O'Reilly Media.

Google Cloud Cookbook: Practical Solutions for Building and Deploying Cloud Services. (2022). O'Reilly Media.

- Hurwitz, J. S., & Kirsch, D. (2020). Cloud Computing for Dummies, 2nd Edition (2e éd.). For Dummies.
- Morris, K. (2021). Infrastructure as Code: Dynamic Systems for the Cloud Age (2e éd.). O'Reilly Media.
- Piper, B., & Clinton, D. (2020). AWS Certified Solutions Architect Study Guide: Associate SAA-C02 Exam (Aws Certified Solutions Architect Official: Associate Exam) (3e éd.). Sybex.

Langue d'enseignement

- Français