

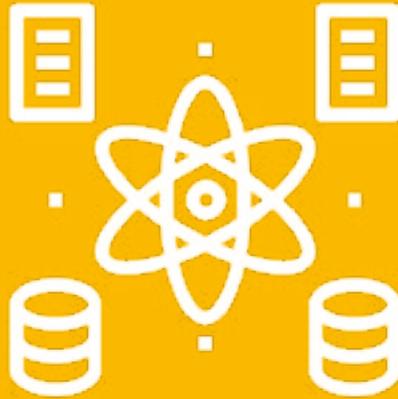
Master 2

Sciences, Technologies, Santé

2020-2021

Mathématiques et applications

Data Science



M2 DS

*Apprentissage
Contrat Pro.*

SOMMAIRE

| | |
|---------------------------------|----|
| Contacts de la formation | 02 |
| Planning de la formation | 03 |
| Présentation de la formation | 05 |
| Volumes horaires et évaluations | 09 |
| Contenu des enseignements | 13 |
| Calendriers détaillés | 25 |

CONTACTS DE LA FORMATION

Assesseure à la Pédagogie

Sandrine TRAVIER

Tél. : 02 41 73 50 01

sandrine.travier@univ-angers.fr

M1 Responsable pédagogique et Président du Jury

Eric DELABAÈRE

Tél. : 02 41 73 54 94

eric.delabaere@univ-angers.fr

Bureau : bâtiment I, porte 214

Responsable pédagogique et Président du Jury M2

Fabien PANLOUP

Tél. : 02 41 73 50 29

fabien.panloup@univ-angers.fr

Bureau : bâtiment I, porte 221

Gestion de la scolarité et des examens

Sandrine HERGUAIS

Tél. : 02 41 73 54 85

sandrine.herguais@univ-angers.fr

Assistante reprise d'études / contrat de professionnalisation

Fatima IKHENOUSSEN

Tél. : 02 44 68 86 96

fatina.ikhenoussen@univ-angers.fr



Scolarité - Examens

Bâtiment A, Rez-de-chaussée, Bureau A006

Horaire d'ouverture

9h00 – 12h30

13h30 – 17h00

Du lundi au vendredi

Fermé le mercredi après-midi

PLANNING DE LA FORMATION

N.B. Calendriers détaillés pour les alternants et non-alternants p. 19/20

Premier semestre

| | |
|--|--|
| Rentrée et début des cours | Jeudi 10 septembre 2020 |
| Campus Day | |
| Forum Ingénierie et Numérique | Novembre 2020 (à préciser) |
| Vacances de la Toussaint | du samedi 24 octobre 2020 au dimanche 1 novembre 2020 |
| Fin des cours du 1 ^{er} semestre | Vendredi 18 décembre 2020 |
| Vacances de Noël | du samedi 19 décembre 2020 au dimanche 3 janvier 2021 |
| Examens 1 ^{er} semestre Session 1 | Du 11 au 15 janvier 2021 |
| Jury 1 ^{er} semestre Session 1 | Mardi 2 février 2021 |

Deuxième semestre

| | |
|---|---|
| Début des cours | Lundi 18 janvier 2021 |
| Fin des cours du 2 ^{ème} semestre | Vendredi 26 février 2021 |
| Examens 2 ^{ème} semestre Session 1 | 1 et 2 mars 2021 |
| Examens 1 ^{er} semestre Session 2 | 2 et 3 mars 2021 |
| Soutenance de Projet Annuel | 4 et 5 mars 2021 non alternants 27 et 28 mai 2021 alternants |
| Soutenance de Stage / Alternance | 2 et 3 septembre 2021 |
| Examens 1 ^{er} semestre Session 1 | Du 11 au 15 janvier 2021 |
| Jury 1 ^{er} semestre Session 2 | 28 septembre 2021 |
| Jury 2 ^{ème} semestre Session 1 | |

Planning susceptible de modifications

PRÉSENTATION
DE LA
FORMATION

Les mathématiques sont depuis longtemps partie prenante du monde socio-économique. «L'Étude de l'impact socio-économique des mathématiques en France » (Mai 2015) montre une contribution des mathématiques à hauteur de 15 % du PIB et 9 % des emplois en France. L'intégration de compétences mathématiques est désormais un choix stratégique du développement d'une entreprise, en France et dans le monde, dans tous les secteurs d'activité et particulièrement ceux à forte croissance.

La data Science désigne une discipline à l'interface entre modélisation mathématique, statistique et informatique, née de la nécessité croissante de traiter et d'exploiter les données du big data. La data science est désormais l'outil essentiel d'aide à la décision dans des domaines d'activités extrêmement variés : banque, finance, assurance, e-commerce et grande distribution, communication et marketing, santé, agro-alimentaire, aéronautique et défense, internet des objets et télécoms, etc..

Le **parcours Data Science (DS)** du master mention Mathématiques et Applications est une formation de niveau un (Bac+5). Il vise à former des cadres data-scientists capables de mettre en œuvre les techniques mathématiques et informatiques inhérentes à ce métier, possédant en outre des compétences métiers propres aux secteurs d'activités visés par le choix d'options.

OBJECTIFS, COMPÉTENCES ET CAPACITÉS VISÉS PAR LA FORMATION

Le master Data Science (DS) est une formation en deux ans qui débouche sur le diplôme national de Master mention Mathématiques et Applications, parcours Data Science. A l'issue de la formation, le ou la diplômé(e) possède :

Des compétences en mathématiques. Le diplômé maîtrise les méthodes de l'analyse statistique et algorithmique de l'information, qu'elles soient classiques (datamining, classification, ..) ou spécifiques au big data (machine learning, ..), ainsi que les méthodes associées d'optimisation. Le diplômé est en capacité de modéliser les processus aléatoires, de les simuler et de confronter le modèle aux données statistiques du problème.

Des compétences en informatique. Le diplômé du master dispose des compétences relatives aux outils classiques de data management et d'analyse des données (R, Python, SQL,..), et plus spécifiques aux données de grandes dimensions et/ou non structurées : plateforme Hadoop, HDFS, NoSQL, etc..

Des aptitudes professionnelles. Le diplômé du master sait définir un projet, ses objectifs et son contexte, les modalités de réalisation et d'évaluation des actions, leurs priorités et leur planning. Il sait utiliser les technologies de l'information et de la communication, effectuer une recherche d'information à partir de sources diverses, organiser une veille technologique. Il a une bonne capacité de communication, une bonne maîtrise de l'anglais scientifique et technique (certification), et sait rédiger clairement une note ou un rapport selon des supports de communication adaptés.

Des compétences métiers. Suivant le choix d'option, le diplômé possède des aptitudes spécifiques au traitement et à l'analyse des données de la biologie (données omiques notamment), ou celles des entreprises dédiées au management du risque ou/et à l'optimisation des actions.

DÉBOUCHÉS VISÉS ET RÉMUNÉRATION DES DIPLÔMÉS

Les diplômés du master Data Science (DS) occupent des emplois de cadre Data scientist (ou Data miner, Data analyst, etc..) couvrant un large spectre de secteurs d'activités, touchant aussi bien au domaine industriel qu'à celui du tertiaire. Sont plus particulièrement visés :

- Option données biologiques : secteurs de la santé, de la pharmacie, la bio-industrie, l'agro-industrie, les laboratoires de type Inserm, Cirad, Institut Pasteur, CEA, etc ..

- Option données numériques : tout secteur tel que sociétés de conseil en gestion des entreprises ou spécialisées en datamining, de la grande distribution, les grandes banques et compagnies d'assurances, etc..

L'enquête interne à 12 mois pour la promotion 2014-2015¹ montre une excellente insertion professionnelle (90 % en emploi CDI ou CDD) pour un salaire moyen brut mensuel de 2 600 Euros (hors primes et 13ème mois).

Selon l'APEC, la rémunération des Data scientists se situe entre 32 et 40 keuros brut/an en début de carrière.

Le master est essentiellement à visée professionnelle. Des possibilités sont cependant offertes aux très bons étudiants d'une poursuite d'étude en doctorat par une thèse, en particulier dans le cadre d'une convention CIFRE avec une entreprise.



¹Enquête pour les diplômés du master Ingénierie Mathématique dont le master DS est une évolution. Exemples de stages ou d'emplois récents : Renault DI-RC, INRA, iAdvize, CPAM Maine et Loire, LNE, Sodi-france-Netapsys, Atmo Poitou Charentes, Air-parif, DAS-MMA, SFR, Ineox, AVISIA, Metro Cash & Carry France, Tronico groupe Alcen, Imbrication, Institut de Cancérologie de l'Ouest, CHU-Angers, Axens IFP Technologies, Coheris, BRED Banque populaire, Enza Zaden, Institut Gustave Roussy, Groupe PSA, CT Ingénierie, Harmonie Mutuelle, Banque de France, VYV Care IT, Capgemini, Sopra-Steria, Sigma Informatique, Socomec SAS,...

PUBLIC VISÉ, PRÉ-REQUIS, CONDITIONS D'ADMISSION, TARIFICATION

Le master Data Science est à capacité limitée et l'admission en première (M1-DS) et deuxième année (M2-DS) du master est sélective. En dehors des qualités académiques, le recrutement accorde un part importante à la motivation des candidats. **La validation du M1-DS entraîne l'admission de droit en M2-DS.**

Formation initiale. Le M1-DS s'adresse en priorité aux étudiants diplômés d'une licence de Mathématiques, d'une licence de Mathématiques et informatique appliquées aux sciences humaines et sociales (parcours Mathématiques appliquées), ou d'un niveau équivalent. Sous réserve d'un niveau suffisant en mathématique, les dossiers d'étudiants ayant eu un parcours plus dirigé vers l'informatique sont considérés avec intérêt. Les droits d'inscription sont fixés par arrêté ministériel (droits ministériels : 243 € en 2019-2020, en M1 et M2).

Formation continue. Les candidatures relevant de la formation continue sont les bienvenues et considérées avec la plus grande attention. Le tarif 2020-2021 est fixé à 6000 € en M2-DS (hors droits ministériels), pour lequel un ensemble de financements est disponible. (Se renseigner auprès de la Direction de la Formation Continue).

Les pré-requis à l'entrée du M1-DS couvrent le programme classique d'une Licence de Ma-thématiques ou tout autre programme européen ou international équivalent. **Outre l'aptitude mathématique, une connaissance et une pratique de base d'outils de programmation sont recommandées** (Python, Scilab, C et/ou R par exemple). L'acquis du certificat C2i-1 d'informatique est apprécié. Le diplôme de master visant des compétences en anglais certifiées de niveau B2 du CECRL, un niveau minimal B1 dans cette langue est recommandé à l'entrée du M1-DS.

L'enseignement se faisant essentiellement en français, **un niveau minimal B2 certifié en français est exigé des étudiants internationaux à l'entrée du M1-DS.**

DES RAISONS DE CHOISIR LE MASTER DATA SCIENCE (DS) DE L'UNIVERSITÉ D'ANGERS

La qualité de la recherche française en mathématiques fondamentales et appliquées est mondialement reconnue : suivant les critères choisis, elle se place du premier au troisième rang mondial. Dans ce cadre, la recherche en mathématique du laboratoire LAREMA de l'Université d'Angers (62 membres dont 40 permanents), associé au CNRS et partenaire du Labex Lebesgue, se distingue par son excellence. Aussi, **les étudiants du Master DS ont l'assurance de bénéficier des connaissances et des innovations les plus récentes**, celles à même d'intéresser les entreprises dans leur stratégie compétitive.

L'Université d'Angers est reconnue par ses résultats en matière de taux de réussite et d'insertion professionnelle de ses 22 700 étudiants : sur ces deux points, les enquêtes ministérielles la situent aux toutes premières places des universités françaises. **Les étudiants du Master DS ont l'assurance d'un accompagnement performant de leurs études jusqu'à leur insertion professionnelle.**

Angers et son agglomération, ville étudiante par excellence (plus de 38 000 étudiants), apparaît dans les enquêtes nationales comme étant l'une des villes les plus attractives dans ce domaine : vitalité de la vie étudiante, qualité et modération des prix du logement étudiant, facilité de déplacement, cadre de vie, etc.. En résumé, **les étudiants du Master DS ont l'assurance d'une qualité de vie et d'étude particulièrement propice à leur réussite et à leur insertion professionnelle.**



VOLUMES HORAIRES ÉVALUATION

SEMESTRE 3 (A)

| SEMESTRE 3 | | | | | | | | | | 30 ECTS | |
|---|--|------|-------|------------------|----|----|----|----------------------------|---------------|---------------------------------|-------------|
| U.E. | Matières | ECTS | Coef. | Volumes horaires | | | | Contrôle des Connaissances | | | |
| | | | | tot. | CM | TD | TP | 1 ^{re} session | | 2 ^e session | Durée exam. |
| | | | | | | | | Assidus | D.A. | | |
| Bloc Mathématiques | | | | | | | | | | | |
| UE1 | Statistique en Grande Dimension et Apprentissage | 6(b) | 6 | 56 | 24 | 16 | 16 | 0.5 CT + 0.5 CC | CT | CT | 3 h 00 |
| UE2 | Introduction à l'Analyse du Signal en Temps Discrets | 4 | 4 | 36 | 16 | 12 | 8 | 0.5 CT + 0.5 CC | CT | CT | 2 h 30 |
| UE3 | Séries Chronologiques et Applications | 6 | 6 | 52 | 24 | 12 | 16 | 0.5 CT + 0.5 CC* | CT | CT | 3 h |
| Bloc Outils numériques et Informatique | | | | | | | | | | | |
| UE4 | Bases de Données Avancées – Big Data Data Visualisation- Exploration de Graphes | 6(b) | 6 | 56 | 20 | 0 | 36 | CC | CT | CT | - - |
| Bloc Culture d'entreprise | | | | | | | | | | | |
| UE5 | Culture d'entreprise | - | - | - | - | - | - | - | - | - | - |
| | Anglais | 2 | 2 | 16 | 0 | 0 | 16 | CC*** | Oral ou Ecrit | Oral ou Ecrit | 1h30 |
| | Communication – Préparation à l'insertion professionnelle | 1 | 1 | 6 | 0 | 0 | 6 | Oral | Oral | Pas de 2 ^{nde} session | - |
| | Conférences professionnelles | 0(c) | 0 | 8 | 8 | 0 | 0 | - | - | - | - |
| Bloc Métiers | | | | | | | | | | | |
| UE6 | Métiers (Options) | 5 | 5 | - | - | - | - | - | - | - | - |
| | Option 1 : Statistique et Science des Données pour la Biologie I | | | 33 | 15 | 8 | 10 | CC | CT | CT | 2 h |
| | Option 2 : Econométrie de l'Evaluation 1 (Commun avec DEG) - Econométrie des choix discrets approfondie | | | 26 | 14 | 12 | | CC | CT | CT | 2 h |

CT : Contrôle Terminal. CC : Contrôle Continu.

* : Épreuve écrite + dossier et soutenance.

CC*** : Oral et écrit (durée 1h30)

CC****: dossier

(a) : **Moyenne supérieure ou égale à 10 au S3** nécessaire pour l'obtention du M2, sauf avis contraire du jury

(b) : **Note plancher de 8**, sauf avis contraire du jury

(c) : Pas d'évaluation mais validation de l'UE par présence : présence obligatoire aux cours (pas de dispense d'assiduité).

Une absence non justifiée entraîne une défaillance à l'UE sauf appréciation contraire du jury.

SEMESTRE 4

| SEMESTRE 4 | | | | | | | | | | 30 ECTS | |
|----------------------------------|---|------|-------|------------------|----|----|----|----------------------------|------|------------------------|-------------|
| U.E. | Matières | ECTS | Coef. | Volumes horaires | | | | Contrôle des Connaissances | | | |
| | | | | tot. | CM | TD | TP | 1 ^{re} session | | 2 ^e session | Durée exam. |
| | | | | | | | | Assidus | D.A. | | |
| Bloc Culture d'entreprise | | | | | | | | | | | |
| UE1 | Culture d'entreprise | - | - | - | - | - | - | - | - | - | - |
| | Conférences professionnelles | 0(a) | 0 | 8 | 8 | 0 | 0 | - | - | - | - |
| | Création d'entreprise | 0(a) | 0 | 3 | 0 | 3 | 0 | - | - | - | - |
| | Projet Annuel Tuteuré | 7 | 7 | - | - | - | - | Oral* | ** | *** | - |
| Bloc Métiers | | | | | | | | | | | |
| UE2 Métiers (Options) | | 5 | 5 | - | - | - | - | - | - | - | - |
| | Option 1 : Statistique et Science des Données pour la Biologie II | | | 40 | 17 | 23 | 0 | CC | ** | *** | 2 h |
| | Option 2 : Econométrie de l'Évaluation 2 (Commun avec DEG) | | | 40 | | | | | | | |
| | - Méthodes d'évaluation appliquées au marché du travail | | | | 14 | 12 | 0 | CC | ** | *** | 2 h |
| | - Econométrie des modèles de durée | | | | 14 | 0 | 0 | CC | ** | *** | 2 h |
| Bloc Stage ou Alternance | | | | | | | | | | | |
| UE3 Stage ou Alternance | | - | - | - | - | - | - | - | - | - | - |
| | Stage ou Alternance | 18 | 18 | - | - | - | - | - | - | - | - |

*Oral : dossier et soutenance.

** : Pas de dispense d'assiduité possible au S4.

*** : Pas de Session 2 au S4

(a) : Pas d'évaluation mais validation par présence : présence obligatoire aux cours (pas de dispense d'assiduité). Une absence non justifiée entraîne une défaillance à l'UE sauf appréciation contraire du jury.

**CONTENU
DES
ENSEIGNEMENTS**

SEMESTRE 3

Les noms des responsables et intervenants sont donnés sous réserve de modifications.

BLOC MATHÉMATIQUES - UE OBLIGATOIRES

UE 1 – Statistique en Grande Dimension et Apprentissage

Statistics for Big Data and Machine Learning

Responsable Fabien Panloup | fabien.panloup@univ-angers.fr
Intervenants Fabien Panloup | fabien.panloup@univ-angers.fr
 Piotr Graczyk | piotr.graczyk@univ-angers.fr

Objectifs (mots clés)

Régression en grande dimension (motivation par l'exemple, théorie du Lasso, seuillage, parcimonie, inégalités classiques, régressions basées sur les composantes principales, PLS). Méthodes de pénalisation (généralisation du Lasso (Elastic-Net, Group-Lasso), sélection de variables, Ridge, etc.). Classification et apprentissage en grande dimension (SVM, Logistic-Lasso, forêts aléatoires, Gradient Boosting, k plus proches voisins, ACP à noyaux), Réseaux de Neurones (Deep Learning), Modèles graphiques. Large mise en pratique avec R ou Python.

Pré-requis

Notions et contenus : Algèbre linéaire et analyse (licence mathématiques L3). Statistique inférentielle (S1-UE3-DS). Optimisation non linéaire (S1-UE2-DS). Datamining et classification (S2-UE1-DS). Manipulation standard de R et de Python.

Compétences et capacités : Maîtriser les compétences enseignées dans les modules de statistique inférentielle et de datamining et classification du M1, particulièrement tout ce qui concerne les modèles de régression linéaire et les méthodes de classification. Maîtriser les bases de l'algèbre linéaire, du calcul matriciel, de l'analyse et de l'optimisation, en particulier : recherche des valeurs propres et singulières, la notion de produit scalaire (sur l'exemple de \mathbb{R}^n et de L^2), les propriétés des espaces de Hilbert de fonctions ou encore l'optimisation sous contraintes. Avoir une connaissance minimale des langages R et Python (syntaxe, manipulation élémentaire, calcul matriciel, gestion des graphiques, etc.).

Connaissances, compétences et capacité exigibles

- Avoir un recul professionnel sur la notion de grande dimension. Comprendre par l'exemple les raisons théoriques mais aussi les raisons techniques qui différencient les méthodes statistiques classiques des méthodes statistiques en grande dimension.
- Comprendre la théorie du Lasso et d'une manière générale la méthodologie consistant à pénaliser une estimation par des contraintes sur les coefficients de la régression. Avoir une vision claire des objectifs visés par les méthodes de pénalisation (estimation ou sélection de variables).
- Savoir réduire la dimension des observations par des méthodes classiques d'analyse des données (par exemple l'ACP) et savoir exploiter cette réduction de dimension pour les régressions en grande dimension ou possédant de la multicollinéarité.
- Savoir ce qui relie et ce qui différencie le Lasso des méthodes de pénalisation basées sur des normes différentes (Elastic-Net, Ridge, etc.). Connaître les points forts et les points faibles de chacune des approches pour en permettre une mise en pratique pertinente sur des jeux de données réelles.

- Connaître les méthodes simples d'apprentissage automatique appliquées à la classification (comme les k plus proches voisins) et les méthodes plus sophistiquées (comme les forêts aléatoires). Avoir une vision d'ensemble de la théorie des SVM appliquée à la classification.
- Être capable de mobiliser et exploiter les méthodes étudiées dans ce module dans des cas pratiques sous R ou Python, en particulier les méthodes de régression et de sélection de variables en grande dimension.
- Connaître les éléments de base des modèles graphiques.

Supports ou références

- *Introduction to High-Dimensional Statistics*. Christophe Giraud. Chapman & Hall/ CRC. 2015
- Trevor Hastie, Robert Tibshirani and Jerome Friedman. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. Springer. 2009.
- Trevor Hastie, Robert Tibshirani, Martin Wainwright. *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press. 2015.

UE 2 – Introduction à l'Analyse du Signal en Temps Discrets

Introduction to Discrete-Time Signal Processing

Responsable Laurent Evain | laurent.evain@univ-angers.fr
 Intervenant Laurent Evain | laurent.evain@univ-angers.fr

Objectifs (mots clés)

Analyse du signal, motivations, exemples. Series de Fourier - Transformation de Fourier sur Z/nZ – Liens entre les deux. Espaces de Hilbert et Energie d'un signal. Filtres FIR (Réponse Impulsionnelle Finie) et IIR (Réponse Impulsion Infinie) – Impulsion – convolution. Transformation en Z – Application à la stabilité des filtres- Problemes d'aliasing

Pré-requis

Notions et contenus : espace euclidien, produit scalaire ; séries de Fourier (licence mathématiques L3).

Compétences et capacités : maîtriser le calcul des coefficients d'une série de Fourier, la décomposition d'une fonction sur une base orthogonale ; connaître et savoir appliquer le théorème de Parseval et les inégalités de Bessel ; maîtriser le langage de programmation Python

Connaissances, compétences et capacité exigibles

- Maîtriser les calculs de transformation de Fourier et de Fourier Inverse.
- Savoir tester la stabilité d'un filtre.
- Savoir concevoir un filtre correspondant à des objectifs donnés par une méthode de fenêtrage ou par l'algorithme de Parks McClellan.
- Savoir utiliser Python et des bibliothèques de type `scipy.signal` pour mettre en pratique la conception des filtres.

Supports ou références

- Prandoni – Vetterli : *Signal Processing for Communications*
- A. V. Oppenheim and R. W. Schaffer, « *Discrete-Time Signal Processing* ». Pearson; 3 edition (2009)
- C. Gasquet, P. Witomski, « *Fourier analysis and applications. Filtering, numerical computation, wavelets.* » *Texts in Applied Mathematics*, 30. Springer-Verlag, New York, 1999

UE 3 – Séries Chronologiques et Applications

Time Series and Applications

Responsable Frédéric Proïa | frederic.proia@univ-angers.fr

Intervenant Frédéric Proïa | frederic.proia@univ-angers.fr

Objectifs (mots clés)

Décomposition des séries chronologiques (modèles additif et multiplicatif, filtrage linéaire). Processus aléatoires discrets (bruit blanc, loi d'un processus, espace L^2). Stationnarité (sens fort, sens faible). Autocorrélations des processus stationnaires (ACF, PACF). Densité spectrale. Processus linéaires (opérateurs chronologiques, décomposition de Wold, bruit d'innovation, causalité). Étude détaillée des processus ARMA. Introduction à la non stationnarité (marche aléatoire, processus ARIMA et SARIMA, élimination de la non stationnarité). Inférence dans les processus ARMA (espérance, autocorrélations et paramètres, propriétés asymptotiques des estimateurs). Application à la prévision (meilleur prédicteur linéaire, lissages exponentiels). Mise en pratique avec R.

Pré-requis

Notions et contenus : Bagage probabiliste de Licence et de Master 1. Algèbre linéaire et analyse (licence mathématiques L3) Statistique inférentielle (S1-UE3-DS). Régression linéaire (S2-UE1-DS). Manipulation standard de R.

Compétences et capacités : Savoir manipuler les opérateurs du calcul des probabilités (espérance, variance, covariance, etc.). Maîtriser les notions de statistique inférentielle du S1 et de régression linéaire du S2, particulièrement ce qui concerne la problématique de l'estimation. Savoir étudier la loi d'un vecteur aléatoire. Maîtriser les bases de l'algèbre linéaire et de l'analyse concernant les espaces vectoriels, les produits scalaires et la projection orthogonale (sur l'exemple de R^n et de L^2). Avoir une connaissance minimale du langage R (syntaxe, manipulation élémentaire, gestion des graphiques, etc...).

Connaissances, compétences et capacité exigibles

- Être capable de reconnaître visuellement une série à composantes additives ou multiplicatives et d'éliminer les tendances par un filtrage approprié.
- Connaître les propriétés de l'espace L^2 des variables aléatoires de carré intégrable, son produit scalaire et la norme associée. En déduire la construction d'une projection orthogonale dans L^2 basée sur les notions de variance et de covariance.
- Savoir étudier la stationnarité d'un processus discret. Comprendre la différence entre stationnarité forte et stationnarité au second ordre.
- Savoir calculer la fonction d'autocorrélation d'un processus stationnaire. Savoir construire son autocorrélation partielle par un algorithme approprié.
- Acquérir une approche théorique solide d'un processus ARMA. Savoir établir des conditions sur ses paramètres permettant d'étudier sa stationnarité, sa causalité, son inversibilité, etc.
- Comprendre la notion de bruit blanc d'innovation. Savoir mettre un processus ARMA sous forme canonique pour l'exprimer en fonction de son bruit blanc d'innovation.
- Savoir estimer l'autocorrélation et l'autocorrélation partielle d'une série chronologique stationnaire, en proposer une modélisation ARMA adaptée par comparaison avec le comportement théorique des autocorrélations.
- Connaître les méthodes d'estimation et les propriétés des estimateurs des paramètres d'un modèle ARMA canonique sous les hypothèses adéquates.
- Savoir étendre la définition des processus ARMA aux processus ARIMA et SARIMA.

- Connaître les principaux tests d'hypothèses adaptés au cadre des processus ARMA (tests de significativité des paramètres, de corrélation ou normalité résiduelle, etc.). Savoir tester la stationnarité d'une série et la stationnariser le cas échéant par l'étude et l'élimination des composantes non stationnaires.
- Être capable de construire le meilleur prédicteur linéaire dans les modèles ARMA, ARIMA ou SARIMA. Connaître les principaux lissages exponentiels.
- Avoir un recul professionnel sur la modélisation d'une série, savoir comparer plusieurs approches sur la base de critères adaptés aux objectifs de l'étude.
- Être capable de mobiliser toutes ces compétences dans des cas pratiques avec R.

BLOC OUTILS NUMÉRIQUES ET INFORMATIQUE - UE OBLIGATOIRE

UE 1 – Outils informatiques

Bases de Données Avancées

Big Data / Advances Databases – Big Data

Responsable François Ducrot | francois.ducrot@univ-angers.fr
 Intervenants François Ducrot | francois.ducrot@univ-angers.fr
 Jacquelin Charbonnel | jacquelin.charbonnel@univ-angers.fr
 Benoît Da Mota | benoit.damota@univ-angers.fr

Objectifs (mots clés)

Technologies du Big Data :

- architecture et composants de la plateforme Hadoop ;
- modes de stockage (NoSQL, HDFS)
- principes de fonctionnement de MapReduce.
- Présentation de technologies spécifiques pour le Big Data
- Gérer les données structurées et non structurées :
- Principes de fonctionnement HDFS
- Importer des données externes vers HDFS
- Réaliser des requêtes SQL avec HIVE
- Utiliser PIG pour traiter la donnée

Pré-requis

Notions et contenus : Bases de données relationnelles (S2-UE4-DS). Programmation (S1-UE5-DS). Statistique (S2-UE1-DS)

Compétences et capacités : Connaissances de base des modèles relationnels. Connaissances en statistique et des langages de programmation.

Connaissances, compétences et capacité exigibles

Comprendre l'écosystème technologique nécessaire pour réaliser un projet Big Data ; acquérir les compétences techniques pour gérer des flux de données complexes, non structurés et massifs ; implémenter des modèles d'analyses statistiques pour répondre aux besoins métiers

Data Visualisation

Data Visualisation

Responsable David Rousseau | david.rousseau@univ-angers.fr
 Intervenant David Rousseau | david.rousseau@univ-angers.fr

Objectifs (mots clés)

Outils logiciels pour la visualisation de données à 1,2, 3, N dimensions, réduction de dimension (ACP, PLS, T-SNE, MDS, LLE, ...), et données orientées (graphes).

Pré-requis

Langage Python/Data Mining

Connaissances, compétences et capacité exigibles

Comprendre l'écosystème technologique nécessaire pour réaliser un projet Big Data ; acquérir les compétences techniques pour gérer des flux de données complexes, non structurés et massifs ; implémenter des modèles d'analyses statistiques pour répondre aux besoins métiers

Supports ou références

— *Python Machine Learning - Second Edition: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow* de Sebastian Raschka (Auteur), Vahid Mirjalili (Auteur)

— *Scikit-learn Cookbook - Second Edition: Over 80 recipes for machine learning in Python with scikit-learn* de Julian Avila (Auteur)

N.B. Évaluation par mini-projet.

Exploitation des données graphes

Mining graph and network data

Responsable [Béatrice Duval | beatrice.duval@univ-angers.fr](mailto:beatrice.duval@univ-angers.fr)

Intervenant [Béatrice Duval | beatrice.duval@univ-angers.fr](mailto:beatrice.duval@univ-angers.fr)

Objectifs (mots clés)

- Notions essentielles sur les graphes (dirigés, non dirigés, graphes bipartite, mesures sur les graphes, pbes essentiels comme plus courts chemins...)
- Illustrations à travers les réseaux sociaux ou les réseaux d'interaction biologiques, outils de visualisation
- Clustering, recherche de communautés dans les graphes

Pré-requis

Langage Python/Data Mining

BLOC CULTURE D'ENTREPRISE – UE OBLIGATOIRES

UE 5 – Culture d'entreprise

Anglais

English

Responsable [Philippe Torrès | philippe.torres@univ-angers.fr](mailto:philippe.torres@univ-angers.fr)

Intervenant [Bruno Menan | bruno.menan@univ-angers.fr](mailto:bruno.menan@univ-angers.fr)

Objectifs (mots clés)

- Permettre aux étudiants de continuer à travailler les cinq compétences en langue - compréhension écrite et orale, expression écrite et orale, et interaction orale - à travers des supports authentiques (articles, documentaires, documents audio et vidéo d'internet, graphiques...) et des activités variées (exercices de compréhension, d'expression écrite, jeux de rôle, débats, présentations orales...) ;

- Etoffer les connaissances lexicales
- Améliorer la prononciation et revoir certains points de langue le cas échéant
- Valider à l'issue du 2ème voire 3ème semestre un niveau B2 du CECRL (Certification)

Pré-requis

Notions et contenus : Les bases de la langue anglaise

Compétences et capacités : Maîtriser le niveau B1 du CECRL (dit « d'utilisateur indépendant »)

Connaissances, compétences et capacité exigibles

On vise, à l'issue du 2ème voire 3ème semestre du Master, le niveau de compétence B2 du Cadre Européen Commun de Référence pour les Langues (CECRL) qui est résumé comme suit : « Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comporte de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités. »

Supports ou références

Les ressources mises à dispositions en autoformations recouvrent entre autres : Numerilang et/ou plateforme LANSAD ; Ressources en lignes (UNT UOH par exemple) ; Logiciels de type Rosetta Stone.

Remarques

Travail additionnel estimé en autoformation : 30h (variable suivant niveau d'entrée).

Communication – Préparation à l'Insertion Professionnelle

Communication – Professional Integration Program

Responsable [Fabien Panloup | fabien.panloup@univ-angers.fr](mailto:fabien.panloup@univ-angers.fr)

Intervenants [Intervenants du SUIO-IP et/ou autres intervenants professionnels](#)

Objectifs

Techniques de communication - Préparation au travail en groupe : Les fonctions de l'animateur; Conduite de réunions - Techniques de rédaction des documents - Techniques de présentation orale - La recherche d'emploi : Lettres de candidature ; Curriculum vitae ; Entretien de candidature

Pré-requis

Compétences de l'UE S2-UE5 communication.

Remarques

Travail additionnel estimé en autoformation : 4h.

Conférences professionnelles

Professional Conferences

Responsable [Fabien Panloup | fabien.panloup@univ-angers.fr](mailto:fabien.panloup@univ-angers.fr)

Intervenants [Intervenants professionnels](#)

Objectifs

Connaître les champs d'applications métiers des connaissances et compétences acquises dans le master. Connaître un ensemble d'entreprises du secteur métier visé. Construire un réseau d'adresses professionnelles

BLOC MÉTIERS – UE OPTIONNELLES (UN CHOIX SUR DEUX)

UE 6 – Métiers

Option 1 : Statistique et Science des Données pour la Biologie I

Responsable Mikael Escobar-Bach | mikael.escobar-bach@univ-angers.fr
Intervenants Jérémy Clotault | jeremmy.clotault@univ-angers.fr
Mikael Escobar-Bach | mikael.escobar-bach@univ-angers.fr
Frédéric Proïa | frederic.proia@univ-angers.fr

Objectifs (mots clés)

Cet enseignement, divisé en trois parties respectivement intitulées « Introduction à l'Analyse de Survie », « Introduction aux plans d'expériences » et « Modèles de reconstruction de généalogie », a pour objectif de présenter quelques thèmes importants de Statistique et Modélisation pour la biologie.

Pré-requis

- Cours de S2-UE6-DS « Introduction à la génomique »
- Cours de Statistique Inférentielle (S1-M1 DS)
- Compétences en R.

Introduction à l'Analyse de Survie

Introduction to Survival Data Analysis

Responsable Mikael Escobar-Bach | mikael.escobar-bach@univ-angers.fr

Le problème des données manquantes est fréquemment rencontré dans un grand nombre de domaines d'applications statistiques. Dans le cadre de la biologie, il n'est pas rare que l'acquisition de données de suivi soit en partie incomplète : pour un essai thérapeutique, le temps de rémission puis de guérison d'un patient peut être trop long et donc jamais observé, on parle alors de données censurées.

Dans ce cours, nous introduirons la méthodologie (fonction de survie, risque instantané, ...) et les outils statistiques (estimateur de Kaplan-Meier, Nelson-Aalen, ...) qui permettent d'appréhender ce genre de problèmes. Une implémentation sur R sera aussi proposée.

Introduction aux plans d'expérience

Introduction to Design of Experiments

Responsable Frédéric Proïa | frederic.proia@univ-angers.fr

Les plans d'expérience forment un outil essentiel en statistique expérimentale, ils permettent d'esquisser les contours d'une méthodologie et d'organiser au mieux la suite des mesures à effectuer afin de répondre à une problématique statistique donnée. Nous décrivons de manière succincte quelques plans d'expérience classiquement utilisés dans les sciences du vivant.

Modèles de reconstruction de généalogie

Probabilistic reconstruction of genealogies

Responsable Jérémy Clotault | jeremmy.clotault@univ-angers.fr

Ce module est dédié à l'étude d'un cas pratique de reconstruction généalogique dans une population végétale. Après quelques concepts génétiques généraux liés aux niveaux de ploïdie, à la production des gamètes et aux croisements, nous proposerons

un modèle permettant de reconstruire, à partir de marqueurs génétiques, un ensemble d'arbres généalogiques probabilisés dans une population végétale soumise à des lois de reproduction spécifiques. Nous étudierons en particulier un modèle 2x (diploïdes), un modèle 4x (tétraploïdes) et, selon le temps disponible, nous réfléchirons à un modèle mixte mettant en interaction des 2x/3x/4x (di/tri/tétraploïdes). Une partie du travail sera dédiée à la programmation effective de la reconstruction généalogique via le logiciel R. La visualisation des résultats sera réalisée à l'aide du package Igraph.

Option 2 : Econométrie de l'Evaluation 1

Econometric Techniques 1

Responsable [Philippe Compaire | philippe.compaire@univ-angers.fr](mailto:philippe.compaire@univ-angers.fr)
Intervenant [Philippe Compaire | philippe.compaire@univ-angers.fr](mailto:philippe.compaire@univ-angers.fr)

Objectifs (mots clés)

Econométrie des choix discrets approfondie.

Connaissances, compétences et capacité exigibles

Logit-Probit-Tobit-Modèles de Heckman. Choix discrets/ Les modèles de choix binaire (probit)/modèle Logit/modèle Logit multinomial/modèle Logit conditionnel/ modèles à choix ordonnés/modèles de comptage/modèles à variables limitées (tronquées)/modèles d'autosélection/Arbres de décision et Nested Logit

Remarques

Ce cours est mutualisé avec les cours de l'unité 2 (S3) du master 2 Economie Appliquée, parcours Ingénierie économique et évaluation des politiques publiques.

SEMESTRE 4

Les noms des responsables et intervenants sont donnés sous réserve de modifications.

BLOC CULTURE D'ENTREPRISE - UE OBLIGATOIRES

UE 1 – Culture d'entreprise

Conférences professionnelles

Professionnal Conferences

Responsable Fabien Panloup | fabien.panloup@univ-angers.fr
Intervenants Intervenants professionnels

Objectifs

Connaître les champs d'applications métiers des connaissances et compétences acquises dans le master. Connaître un ensemble d'entreprises du secteur métier visé. Construire un réseau d'adresses professionnelles

Création d'Entreprise

Entrepreneurship Training

Responsable Fabien Panloup | fabien.panloup@univ-angers.fr
Intervenants Intervenants du SUIO-IP et/ou autres intervenants professionnels

Projet Annuel Tuteuré

Annual Training Project

Responsable Fabien Panloup | fabien.panloup@univ-angers.fr
Intervenants Variable suivant le sujet proposé

Objectifs

Deux cas sont à distinguer :

Formation Initiale : les étudiants réalisent en groupe un projet annuel entre septembre et février avec une soutenance à mi-projet, sous l'encadrement d'un tuteur enseignant. La réalisation du projet s'efforcera de calquer le déroulement de la gestion de projets dans le monde socio-économique. Dans une première phase, après avoir circonscrit le problème, un cahier des charges sera réalisé et un calendrier sera défini pour la réalisation. La seconde phase du projet consistera au traitement du problème en utilisant les concepts, méthodes et outils appropriés et satisfaisants au cahier des charges. Le projet annuel met en oeuvre l'ensemble des compétences théoriques et pratiques du master, sur un thème en lien avec l'option métier.

Alternants : les étudiants alternants réalisent un projet défini en partenariat avec l'entreprise d'accueil.

Supports ou références

Variable suivant sujet

BLOC MÉTIERS – UE OPTIONNELLES (UN CHOIX SUR DEUX)

UE 2 – Métiers

Option 1 : Statistique et Sciences des Données pour la Biologie II

Responsable [Claudine Landès](mailto:claudine.landes@univ-angers.fr) | claudine.landes@univ-angers.fr

Cet enseignement est divisé en deux parties : la première intitulée « Bioimage processing » est consacrée au traitement d'images tandis que la seconde intitulée « Inférence de réseaux et transcriptomique » a pour objectif de renforcer l'analyse de données génomiques et d'introduire les réseaux biologiques. Ci-dessous, un descriptif de chacune des parties :

Pré-requis

- Contenu du cours S2-M1 DS « Introduction à la génomique ».
- Contenus des cours de S3-M2 DS « Statistique en Grande Dimension et Apprentissage » et « Bases de Données Avancées-Big Data ».

Traitement d'Images pour la Biologie

Bioimage Processing

Intervenant [David Rousseau](mailto:david.rousseau@univ-angers.fr) | david.rousseau@univ-angers.fr

Projet Annuel Tuteuré

Annual Training Project

Responsable [Fabien Panloup](mailto:fabien.panloup@univ-angers.fr) | fabien.panloup@univ-angers.fr

Intervenants [Variable suivant le sujet proposé](#)

Objectifs

Cet enseignement a pour objectif d'introduire les concepts principaux du traitement d'images pour les mettre en oeuvre via des outils logiciels libres standards (Scikit image Python, imageJ, Illastik, Knime, ...) sur des questions biologiques variées (invitro, invivo, exvivo, macroscopiques-microscopiques, biomédicales, sciences des plantes, ...) . A l'issue du cours, les étudiants devront être capables de construire un « pipeline » de traitement d'images pour réaliser de l'apprentissage machine basé expert ou par réseau de neurones profonds.

Inférence de Réseaux et Transcriptomique

Inference for Biological Networks and Transcriptomic data analysis

Intervenants [Sébastien Aubourg](mailto:sebastien.aubourg@univ-angers.fr) | sebastien.aubourg@univ-angers.fr

[Béatrice Duval](mailto:beatrice.duval@univ-angers.fr) | beatrice.duval@univ-angers.fr

[Olivier Goudet](mailto:olivier.goudet@univ-angers.fr) | olivier.goudet@univ-angers.fr

[Claudine Landès](mailto:claudine.landes@univ-angers.fr) | claudine.landes@univ-angers.fr

Objectifs

Introduction aux réseaux biologiques / Méthodes d'inférences de réseaux / Causalité / Méthode de validation.

Cet enseignement sera dispensé sous forme de CM et de TP de mise en pratique sur des données réelles. La réalisation d'un projet de traitement de données transcriptomiques et d'inférence de réseaux sera proposée à l'issue du cours.

Option 2 : Econométrie de l'Evaluation 2

Econometric Techniques 2

Responsable Philippe Compaire | philippe.compaire@univ-angers.fr
Intervenants Christophe Daniel | christophe.daniel@univ-angers.fr
 Jésus Nze Obame | jesus.nzeobame@univ-angers.fr

Objectifs

- 1- Méthodes d'évaluation appliquées au marché du travail
- 2- Econométrie des modèles de durée

Connaissances, compétences et capacité exigibles

- 1- Biais d'endogénéité, de sélection, d'hétérogénéité, Méthodes des variables instrumentales, de l'Inverse du Ratio de Mill, estimation par quantiles, doubles différences, score de propension, appariement.
- 2- Données de survie, fonction de hasard, fonction de survie, Kaplan-Meier, méthode actuarielle, risques proportionnels constants, modèles de Cox, estimations et lois paramétriques de durée

Remarques

Ce cours est mutualisé avec les cours de l'unité 1 (S4) du master 2 Economie Appliquée, parcours Ingénierie économique et évaluation des politiques publiques.

BLOC STAGE OU ALTERNANCE

UE 3 Stage ou Alternance

Stage de 5 à 6 mois en entreprise ou alternance sur l'année (voir calendrier p. 19/20)

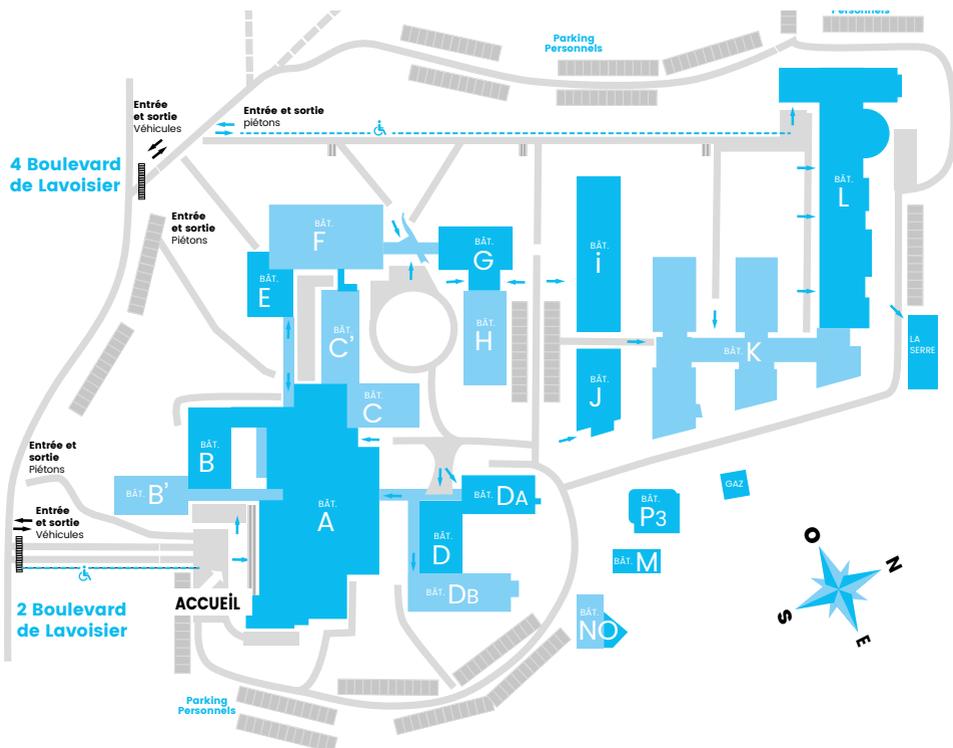
CALENDRIERS

CALENDRIER DÉTAILLÉ 2020-2021 FORMATION INITIALE/ALTERNANTS

| Semaines | # | Début | Fin | Formation initiale | Alternance |
|--------------|------|------------|------------|--------------------------|--------------------------|
| 36-37 | 1,5 | 1/09/2020 | 09/09/2020 | | Entreprise |
| 37-43 | 6,5 | 11/09/2020 | 23/10/2020 | Enseignements S3 | Enseignements S3 |
| 44-46 | 3 | 26/10/2020 | 13/11/2020 | Projet Annuel | Entreprise |
| 47-51 | 5 | 16/11/2020 | 18/12/2020 | Enseignements S3 | Enseignements S3 |
| 51-52 | 2 | 21/12/2020 | 31/12/2020 | Vacances | Entreprise |
| 1 | 1 | 04/01/2021 | 08/01/2021 | Projet Annuel | Entreprise |
| 2 | 1 | 11/01/2021 | 15/01/2021 | Examens S3 | Examens S3 |
| 3-4 | 2 | 13/01/2021 | 07/02/2021 | Enseignements S4 | Enseignements S4 |
| 5-6 | 2 | 1/02/2021 | 12/02/2021 | Projet Annuel | Entreprise |
| 7-8-9 | 2,5 | 13/01/2021 | 03/03/2021 | Enseignements S4 | Enseignements S4 |
| 9 | 0,5 | 4/03/2021 | 5/03/2021 | Soutenance Projet Annuel | Entreprise |
| 10-18 | 9 | 8/03/2021 | 7/05/2021 | Stage | Entreprise |
| 19-21 | 3 | 10/05/2021 | 28/05/2021 | Stage | Projet Annuel+Soutenance |
| 22-35 | 12,5 | 31/05/2021 | 31/08/2021 | Stage | Entreprise |
| 35 | 0,5 | 01/09/2021 | 03/09/2021 | Soutenance Stage | Soutenance Stage |

CALENDRIER ALTERNANCE (CONTRAT PRO OU APPRENTISSAGE)





- A** Administration | Scolarité | Enseignement (Amphi A à E)
- B** Biologie végétale | Physiologie végétale | Travaux pratiques biologie
- B'** Travaux pratiques biologie
- C** Travaux pratiques chimie
- C'** Département de Géologie | Recherche environnement (LETG -LEESA) | Recherche géologie (LPGN-BIAF)
- D** Travaux pratiques physique
- Da** Enseignement | Travaux pratiques physique
- Db** Département de Physique | Recherche physique (LPHIA)
- E** Travaux pratiques biologie
- F** Département de Biologie | Recherche neurophysiologie (SIFCIR) | Travaux pratiques biologie, géologie
- GH** Département informatique | Recherche informatique (LERIA) | Travaux pratiques géologie
- i** Département Mathématiques | Recherche Mathématiques (LAREMA)
- J** Chimie enseignement | Travaux pratiques
- K** Département de Chimie | Recherche Chimie (MOLTECH Anjou)
- L** Espace multimédia | Enseignement (Amphi L001 à L006) | Salle d'examen rez-de-jardin



**FACULTÉ
DES SCIENCES**
UNIVERSITÉ D'ANGERS

2, Boulevard Lavoisier
49045 ANGERS CEDEX 01
T.0241735353
www.univ-angers.fr