

PÉRIODE D'ACCRÉDITATION : 2016 / 2021

UNIVERSITÉ PAUL SABATIER

SYLLABUS MASTER

Mention Mathématiques et applications

M2 statistiques et Informatique décisionnelle

<http://www.fsi.univ-tlse3.fr/>
[http://departement-math.univ-tlse3.fr/
master-mention-mathematiques-et-applications-620690.kjsp](http://departement-math.univ-tlse3.fr/master-mention-mathematiques-et-applications-620690.kjsp)

2019 / 2020

28 OCTOBRE 2019

SOMMAIRE

PRÉSENTATION	3
PRÉSENTATION DE LA MENTION ET DU PARCOURS	3
Mention Mathématiques et applications	3
Parcours	3
PRÉSENTATION DE L'ANNÉE DE M2 statistiques et Informatique décisionnelle	3
Liste des formations donnant accès de droit :	4
RUBRIQUE CONTACTS	5
CONTACTS PARCOURS	5
CONTACTS MENTION	5
CONTACTS DÉPARTEMENT : FSI.Math	5
Tableau Synthétique des UE de la formation	6
LISTE DES UE	7
GLOSSAIRE	25
TERMES GÉNÉRAUX	25
TERMES ASSOCIÉS AUX DIPLOMES	25
TERMES ASSOCIÉS AUX ENSEIGNEMENTS	25

PRÉSENTATION

PRÉSENTATION DE LA MENTION ET DU PARCOURS

MENTION MATHÉMATIQUES ET APPLICATIONS

L'objectif du master mention Mathématiques et Applications est de former des mathématiciens pouvant travailler dans les métiers liés à l'**ingénierie**(parcours MApl3, SID, RO, SE, RI), à la **recherche**(parcours RI, RO, MApl3) et à l'**enseignement**(parcours ES),

Les métiers de l'ingénierie sont typiquement chefs de projets, chargés d'études, ingénieurs et chercheurs dans des secteurs d'activités tels que l'industrie, les services, le marketing.

Les métiers de l'enseignement concernent des postes de professeur de mathématiques en lycée, à l'université en passant par les classes préparatoires.

La recherche peut-être de nature académique, théorique et/ou appliquée, ou être tournée vers l'innovation et le développement dans le secteur privé.

Que ce soit pour les métiers de l'ingénierie, de l'enseignement ou de la recherche le nombre d'étudiants formés aux mathématiques en France est très inférieur au nombre de postes à pourvoir. De ce fait, l'insertion des étudiants titulaires d'un master en Mathématiques est excellente

PARCOURS

Le parcours SID-Big Data (Statistique et Informatique Décisionnelle), labellisé CMI (Cursus de Master en Ingénierie) est une formation résolument tournée vers les métiers de l'ingénierie des données, qui en aborde tous les aspects : de la collecte à l'exploitation statistique et au machine learning en passant par le stockage et la gestion de bases de données massives et distribuées de type big data.

Il permet d'acquérir une double compétence en Statistique et en Informatique décisionnelle, complétée par une connaissance de l'entreprise et une pratique professionnelle acquise au travers des projets et de plusieurs stages. Cette double compétence, à l'ère du Big Data, est très appréciée sur le marché de l'emploi.

Les débouchés de la formation SID-Big Data (Licence et Master) se trouvent dans tous les secteurs où la mise en place, l'administration, l'exploitation et la valorisation de l'information sont primordiales pour la prise de décision.

PRÉSENTATION DE L'ANNÉE DE M2 STATISTIQUES ET INFORMATIQUE DÉCISIONNELLE

Une fois acquis les concepts fondamentaux en Mathématiques/Statistique et Informatique/Base de données en L3 SID, les enseignements du Master SID-Big Data sont axés sur le perfectionnement et la spécialisation de méthodes plus avancées, avec un juste équilibre entre informatique/bases de données et mathématiques/statistique. En Master 1, les méthodes classiques dans le domaine des sciences des données sont approfondies et des méthodes plus novatrices, orientée vers la big data sont introduites. En Master 2, toutes ces compétences sont renforcées pour que les étudiants puissent maîtriser de bout en bout les différents aspects de leur futur métier dans le domaine des sciences des données.

Les principales thématiques abordées en Master SID-Big Data :

- Statistique exploratoire
- Modélisation aléatoire
- Séries chronologiques, sondages
- Apprentissage / Big data
- Exploitation et analyse de données texte, audio, vidéo
- Data mining / Text mining
- Optimisation numérique

- Base de données avancées
- Business intelligence

Au cours des deux années du Master SID-Big Data, en parallèle des enseignements classiques (cours, TD et TP), de nombreux projets sont mis en place, balayant l'ensemble des compétences enseignées.

En M1, un stage de 3-4 mois est obligatoire au 2nd semestre.

Les étudiants peuvent suivre le M2 selon deux formules :

- en formation initiale, sur un mode d'organisation classique : cours, TD, TP, projets, et enfin un stage de 5-6 mois couvrant tout le 2nd semestre.
- en alternance, avec des périodes (de 2 à 4 semaines) à l'université (avec les étudiants en formation initiale) ou en entreprise.

LISTE DES FORMATIONS DONNANT ACCÈS DE DROIT :

M1 STATISTIQUE ET INFORMATIQUE DECISIONNELLE (EMMABE)

Pour les étudiants ayant suivi une autre formation que l'année précédente du parcours, l'accès est sur dossier. Il est très fortement conseillé de se rapprocher du responsable de la formation envisagée pour en connaître les modalités d'accès.

RUBRIQUE CONTACTS

CONTACTS PARCOURS

RESPONSABLE M2 STATISTIQUES ET INFORMATIQUE DÉCISIONNELLE

HUBERT Gilles
Email : hubert@irit.fr

CHOUQUET Cécile
Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

SECRÉTAIRE PÉDAGOGIQUE

NICOLAS Clement
Email : clement.nicolas2@univ-tlse3.fr

CONTACTS MENTION

RESPONSABLE DE MENTION MATHÉMATIQUES ET APPLICATIONS

COSTANTINO Francesco
Email : Francesco.Costantino@math.univ-toulouse.fr

CONTACTS DÉPARTEMENT: FSI.MATH

DIRECTEUR DU DÉPARTEMENT

BUFF Xavier
Email : xavier.buff@univ-tlse3.fr

Téléphone : 5 76 64

SECRETARIAT DU DÉPARTEMENT

RODRIGUES Manuella
Email : manuella.rodrigues@univ-tlse3.fr

Téléphone : 05 61 55 73 54

Université Paul Sabatier
1TP1, bureau B13
118 route de Narbonne
31062 TOULOUSE cedex 9

TABLEAU SYNTHÉTIQUE DES UE DE LA FORMATION

9

page	Code	Intitulé UE	ECTS	Obligatoire Facultatif	Cours	TD	TP	Projet	Stage
Premier semestre									
8	EIMAB3AM	MODÉLISATION STATISTIQUE AVANCÉE, BE	3	O					
9		EIMAB3A1 Modélisation Statistique avancée			6	10	4		
		EIMAB3A2 BE "Statistique & Santé" ou "Statistique & industrie"						50	
10	EIMAB3BM	STATISTIQUE POUR L'INDUSTRIE	3	O	12	12	6		
11	EIMAB3CM	APPRENTISSAGE POUR LES DONNÉES MASSIVES	3	O	12	6	12		
12	EIMAB3DM	BASES DE DONNÉES	6	O					
13		EIMAB3D1 BD massivement parallèles, cloud computing			8	12			
14		EIMAB3D2 Business Intelligence			8	6	6		
		EIMAB3D3 Bases de données semi et non structurées			8	6	6		
15	EIMAB3EM	ÉVALUATION ET CERTIFICATION LOGICIEL	3	O					
16		EIMAB3E1 Evaluation et certification logiciel			12	18			
		EIMAB3E2 Evaluation et certification logiciel (BE)						50	
17	EIMAB3FM	COMMUNICATION ET MANAGEMENT EN ENTREPRISE	3	O	12	18			
18	EIMAB3HM	ANALYSE ET EXPLOITATION DE DONNÉES	6	O					
19		EIMAB3H1 Analyse et exploitation de données			16	8	16		
20		EIMAB3H2 Analyse et exploitation de données (TER/Challenge)						100	
20	EIMAB3VM	ANGLAIS	3	O		24			
Second semestre									
21	EIMAB4AM	NOSQL ET EXPERIENCES INDUSTRIELLES EN SCIENCES DES DONNEES	3	O	30				
22	EIMAB4BM	PROJET / CHALLENGE OPEN-DATA OU BIG-DATA	6	O				150	
Choisir 1 UE parmi les 2 UE suivantes :									
23	EIMAB4CM	STAGE	21	O					5
24	EIMAB4DM	STAGE	21	O					5

LISTE DES UE

UE	MODÉLISATION STATISTIQUE AVANCÉE, BE	3 ECTS	1^{er} semestre
Sous UE	Modélisation Statistique avancée		
EIMAB3A1	Cours : 6h , TD : 10h , TP : 4h		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

OBJECTIFS D'APPRENTISSAGE

Ce cours vient en complément des enseignements sur le modèle linéaire sur des situations plus complexes, telles que les données répétées, les réponses multidimensionnelles ou les modèles mixtes.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Une première partie de cet enseignement est basée principalement sur les méthodes de modélisation déjà étudiées en L3 et M1 avec quelques extensions.

La seconde partie constitue le cœur de cet enseignement avec des nouvelles méthodes de modélisation permettant de prendre en compte des données particulières :

- Analyse de variance multidimensionnelle (MANOVA) pour modéliser une variable multidimensionnelle
- Modèles à effets aléatoires et modèles mixtes pour prendre en compte des facteurs à effet aléatoire
- Modèles pour données répétées pour modéliser une variable réponse mesurée à plusieurs reprises au cours du temps pour un même individu.

PRÉ-REQUIS

Modèles linéaires gaussiens et généralisés

RÉFÉRENCES BIBLIOGRAPHIQUES

A.C. Rencher, "Methods of multivariate analysis", Wiley, 1995.

T.W. Anderson, "An introduction to multivariate statistical analysis", Wiley, 2003.

MOTS-CLÉS

Modélisation statistique, données répétées, réponse multidimensionnelle, modèle mixte, effet aléatoire

UE	MODÉLISATION STATISTIQUE AVANCÉE, BE	3 ECTS	1^{er} semestre
Sous UE	BE "Statistique & Santé" ou "Statistique & industrie"		
EIMAB3A2	Projet : 50h		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

OBJECTIFS D'APPRENTISSAGE

L'objectif de ce module est de mettre en œuvre une étude statistique sur des données réelles issues du milieu industriel ou du secteur de la santé.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Cet enseignement s'articule autour d'un projet d'analyse statistique sur des données réelles. Il permet aux étudiants de revoir l'ensemble des méthodes statistiques vues pendant leur cursus, mais aussi d'appréhender de nouveaux outils mathématiques et statistiques spécifiquement adaptées aux données étudiées. L'objectif est de montrer comment faire face aux difficultés inhérentes aux données réelles, par exemple les données manquantes. Dans le cadre d'un projet sur données biomédicales, nous pourrions présenter en particulier les méthodes d'analyse de survie pour analyser les données censurées (Méthode actuarielle, méthode de Kaplan-Meier, Modèle de Cox).

PRÉ-REQUIS

Statistique descriptive et inférentielle, modèles linéaires, Analyse des données, apprentissage

RÉFÉRENCES BIBLIOGRAPHIQUES

Statistical Modeling for Biomedical Researchers : A Simple Introduction to the Analysis of Complex Data, William D. Dupont

Statistical Analysis with Missing Data. R.J.A.Little et D.B.Rubin

MOTS-CLÉS

Etude statistique, données réelles, analyse, modélisation

UE	STATISTIQUE POUR L'INDUSTRIE	3 ECTS	1^{er} semestre
EIMAB3BM	Cours : 12h , TD : 12h , TP : 6h		

ENSEIGNANT(E) RESPONSABLE

PAUWELS Edouard

Email : edouard.pauwels@irit.fr

OBJECTIFS D'APPRENTISSAGE

Ce cours donne un large panorama sur la maîtrise statistique des procédés (MSP). Il s'agit de techniques descriptives et inférentielles pour le contrôle statistique des processus.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Rappels et compléments sur la distribution gaussienne. Méthodes d'estimation des paramètres.

Capabilités. Cartes de contrôles et applications.

Plans d'échantillonnage.

PRÉ-REQUIS

Bases en probabilités et statistiques, loi normale, estimateur, test.

RÉFÉRENCES BIBLIOGRAPHIQUES

« MSP : Principes et cas industriels », Mercier S. et Bergeret F.

MOTS-CLÉS

Capabilité, cartes de contrôle, plan d'échantillonnage.

UE	APPRENTISSAGE POUR LES DONNÉES MASSIVES	3 ECTS	1^{er} semestre
EIMAB3CM	Cours : 12h , TD : 6h , TP : 12h		

ENSEIGNANT(E) RESPONSABLE

PAUWELS Edouard

Email : edouard.pauwels@irit.fr

OBJECTIFS D'APPRENTISSAGE

L'objectif de cette UE est d'étendre les compétences acquises auparavant sur l'apprentissage statistique à des configurations où les données sont massives : nous verrons quels nouveaux problèmes cela pose, quelles méthodes il est possible d'adapter, et quels nouveaux algorithmes on peut utiliser.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- Rappels et compléments sur la sélection de modèle et la pénalisation L1 (algorithme LASSO)
- Apprentissage distribué : paradigme map-reduce, environnements distribués et passage à l'échelle (notions)
- Tests multiples : corrections de Bonferroni, de Holm-Bonferroni, False Discovery Rate (FDR)
- Méthodes d'agrégation pour l'apprentissage, apprentissage dans les flux de données
- Méthodes d'optimisation pour l'apprentissage dans les données massives, deep learning, gradient stochastique et applications
- Recommandation de contenu et complétion de matrices : modèles de bandits et factorisation non-négative de matrices
- Introduction au traitement statistique de la confidentialité

PRÉ-REQUIS

Machine Learning : théorie et algorithmes classiques, Modélisation aléatoire (statistique et probabilité), Notions d'algorithmique distribuée

RÉFÉRENCES BIBLIOGRAPHIQUES

Amini, apprentissage machine : de la théorie à la pratique

Cesa-Bianchi & Lugosi, Prediction, Learning and Games

MOTS-CLÉS

apprentissage statistique, mégadonnées, flux, analyse de données

UE	BASES DE DONNÉES	6 ECTS	1^{er} semestre
Sous UE	BD massivement parallèles, cloud computing		
EIMAB3D1	Cours : 8h , TD : 12h		

ENSEIGNANT(E) RESPONSABLE

HAMEURLAIN Abdelkader
Email : hameurlain@irit.fr

OBJECTIFS D'APPRENTISSAGE

Comprendre les principaux problèmes posés et les méthodes proposées dans la conception des SGBD parallèles et les systèmes clouds.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

1. Objectifs des BD parallèles
2. Méthodes de répartition de données
3. Stratégies de parallélisation de requêtes décisionnelles
4. Introduction aux traitements de données dans les systèmes clouds
5. Évaluation et optimisation élastique de requêtes décisionnelles
6. Étude comparative entre les SGBD parallèles et les systèmes clouds

PRÉ-REQUIS

Systèmes de BD relationnels, Evaluation et optimisation de requêtes, Systèmes de BD distribués

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Principles of Distributed Database Systems, T. Oszu P. Valduriez, 3rd edition, 2011
2. Traitement parallèle dans les bases de données relationnelles : concepts, méthodes et applications, A. Hameurlain, P. Bazex, F. Morvan, 1996.

MOTS-CLÉS

Répartition des des données, Parallélisme, Parallélisation de requêtes, Systèmes Clouds, Scalabilité, Elasticité

UE	BASES DE DONNÉES	6 ECTS	1^{er} semestre
Sous UE	Business Intelligence		
EIMAB3D2	Cours : 8h , TD : 6h , TP : 6h		

ENSEIGNANT(E) RESPONSABLE

HUBERT Gilles
 Email : hubert@irit.fr

OBJECTIFS D'APPRENTISSAGE

Connaître les outils de Business intelligence.
 Maîtriser l'utilisation d'un outil de business intelligence.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Business Intelligence d'une manière globale
 ERP (Editeurs et Open source)
 Business Object
 OLAP
 Outils de reporting
 Big Data
 ECM (Entreprise Content Management)
 MDM (Master Data Management)
 CRM (Customer Relationship Management)
 SOA (Service Oriented Architecture)...
 Utilisation avancée de Business Object.

PRÉ-REQUIS

Bases de données relationnelles, SQL, Entrepôts de données, principes et outils ETL

RÉFÉRENCES BIBLIOGRAPHIQUES

Successful Business Intelligence, Second Edition : Unlock the Value of BI & Big Data 2nd Edition, C. Howson, Mc Graw Hill Education
 SAP BusinessObjects BI 4.0 : The Complete Reference, Cindi Howson, E. Newbould, C. Duey, Osborne/McGraw-Hill

MOTS-CLÉS

Business Intelligence, OLAP, Business Object, Big data

UE	BASES DE DONNÉES	6 ECTS	1^{er} semestre
Sous UE	Bases de données semi et non structurées		
EIMAB3D3	Cours : 8h , TD : 6h , TP : 6h		

ENSEIGNANT(E) RESPONSABLE

BAHSOUN Wahiba
 Email : wbahsoun@irit.fr

Téléphone : 05 61 55 69 45

OBJECTIFS D'APPRENTISSAGE

Apprendre à manipuler le modèle de données XQUERY.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

XQUERY

- Structure d'arbre d'un document XML
- Concepts fondamentaux du langage de requêtes Xpath : expressions, sélecteurs de nœuds, axe de recherche, filtres, fonctions et opérations.
- Modèle de données Xquery : Forêts, Collections d'arbres...
- Projection d'arbres sur des sous-arbres, sélection d'arbres et des sous-arbres.
- Langage Xquery par l'exemple

PRÉ-REQUIS

Bases de données relationnelles, algorithmique et programmation.

RÉFÉRENCES BIBLIOGRAPHIQUES

XQuery 2e, Priscilla Walmsley, O'Reilly
 XQuery pour interroger des données XML, Jacques Lemaître, Emmanuel Bruno, Ellipses

MOTS-CLÉS

XQuery, XML

UE	ÉVALUATION ET CERTIFICATION LOGICIEL	3 ECTS	1^{er} semestre
Sous UE	Evaluation et certification logiciel		
EIMAB3E1	Cours : 12h , TD : 18h		

ENSEIGNANT(E) RESPONSABLE

BAHSOUN Wahiba

Email : wbahsoun@irit.fr

Téléphone : 05 61 55 69 45

OBJECTIFS D'APPRENTISSAGE

Donner un avis d'expert sur les critères d'Evaluation de Logiciel évoqués dans le cursus de la formation.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- Approche de la qualité du logiciel à travers le processus de développement (Modèle CMM, Standards et Normes de développements...) et le Produit (Norme ISO9126, Techniques de Tests,...)
- Présentation des normes (DO2167, DO178), outils et techniques appliquées dans le domaine aéronautique, spatial et des systèmes "Safety of Life".
- Introduction de la notion de certification Logiciel : Processus, Outils Case, Traçabilité des exigences qualité mises en œuvre dans un contexte de développement.
- Diagnostic Qualité, arbre de défaillance, Méthode d'Evaluation de l'arbre qualité.

PRÉ-REQUIS

Maitriser les concepts des différents modèles de développement.

RÉFÉRENCES BIBLIOGRAPHIQUES

Software Engineering in Higher Education Editors G.King, C.A. Brebia, M.Ross, G. Staples

MOTS-CLÉS

outils case, cycle de vie, exigences

UE	ÉVALUATION ET CERTIFICATION LOGICIEL	3 ECTS	1^{er} semestre
Sous UE	Evaluation et certification logiciel (BE)		
EIMAB3E2	Projet : 50h		

ENSEIGNANT(E) RESPONSABLE

BAHSOUN Wahiba

Email : wbahsoun@irit.fr

Téléphone : 05 61 55 69 45

OBJECTIFS D'APPRENTISSAGE

Je sais mettre en oeuvre mes connaissances et compétences en Evaluation et certification logicielle dans un cas d'étude concret.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Bureau d'étude/Cas d'étude concret relatif à l'évaluation et la certification logicielle.

PRÉ-REQUIS

Maitriser les concepts des différents modèles de développement.

RÉFÉRENCES BIBLIOGRAPHIQUES

Software Engineering in Higher Education Editors G.King, C.A. Brebia, M.Ross, G. Staples

MOTS-CLÉS

outils case, cycle de vie, exigences

UE	COMMUNICATION ET MANAGEMENT EN ENTREPRISE	3 ECTS	1^{er} semestre
EIMAB3FM	Cours : 12h , TD : 18h		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

HUBERT Gilles

Email : hubert@irit.fr

OBJECTIFS D'APPRENTISSAGE

Ce module a deux objectifs :

- fournir aux étudiants les concepts de base du cycle de décision dans les entreprises
- savoir préparer et mener un entretien d'embauche professionnel.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Cycle de décision en entreprise :

En cycle de décision nous développons les deux principales méthodes de prise de décision et présentons l'approche systémique de l'entreprise. Nous développons les trois niveaux de la prise de décision (stratégique, tactique, opérationnel). L'intelligence économique est également présentée à la fois comme un mode de gouvernance et en tant que science. Les trois axes de l'intelligence économique sont détaillés, avec un accent particulier mis sur la veille stratégique.

Techniques de communication :

Ce cours aborde l'entretien d'embauche en face à face : l'importance de la préparation et du non verbal, les techniques professionnelles relevant de l'entretien, les erreurs à éviter, les outils pour suivre sa candidature.

PRÉ-REQUIS

Bonne connaissance de ce que sont les systèmes d'information.

RÉFÉRENCES BIBLIOGRAPHIQUES

M Salles. Décision et système d'information.

JR Alcaras. Les conceptions de la décision en sciences économiques : vers une approche ingénieriale ?

MOTS-CLÉS

Cycle, prise de décision, Intelligence économique, Veille stratégique

UE	ANALYSE ET EXPLOITATION DE DONNÉES	6 ECTS	1^{er} semestre
Sous UE	Analyse et exploitation de données		
EIMAB3H1	Cours : 16h , TD : 8h , TP : 16h		

ENSEIGNANT(E) RESPONSABLE

FARINAS Jérôme

Email : jerome.farinas@irit.fr

Téléphone : 0561558343

LECHANI-TAMINE Lynda

Email : Lynda.Tamine-Lechani@irit.fr

Téléphone : 0561533881

PINQUIER Julien

Email : pinquier@irit.fr

Téléphone : 7434

OBJECTIFS D'APPRENTISSAGE

Être capable de traiter des informations brutes de type texte, audio, vidéo en vue d'en dégager des informations exploitables par les décideurs.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- I. Objectifs, enjeux et applications
- II. Exploitation de textes et applications
 1. Catégorisation de textes : Méthodes génératives (Naive Bayes, Bernoulli), Méthodes vectorielles (Rocchio, KNN), Méthodes basées sur l'apprentissage (SVM)
 2. Classifications de textes : Méthode k-means, Méthode EM
 3. Résumé automatique
 4. Cadre d'application : analyse et résumé d'opinions (en lien avec la partie projet/challenge)
- III. Exploitation de contenus sonores et applications
 1. Classification de sons (phonèmes, locuteurs) : Méthodes GMM, Méthodes HMM
 2. Recherche de mots clefs
 3. Transcription automatique de la parole
 4. Indexation audio
- IV. Exploitation de contenus visuels et applications
 1. Analyse de la texture
 2. Analyse de formes et de contours
 3. Classification et recherche d'images par le contenu : vers l'indexation vidéo

PRÉ-REQUIS

Extraction d'information (S7)

Méthode de catégorisation/classification (K-means, KNN, SVM, EM)

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Modèles statistiques pour l'accès à l'information textuelle, Eric Gaussier, François Yvon
2. Multimodal Processing and Interaction. Audio, Video, Text, Springer, Petros Maragos
3. Structured Analysis Design Technic, Michel Galinier

MOTS-CLÉS

modélisation automatique, texte, audio, video, indexation par le contenu

UE	ANALYSE ET EXPLOITATION DE DONNÉES	6 ECTS	1^{er} semestre
Sous UE	Analyse et exploitation de données (TER/Challenge)		
EIMAB3H2	Projet : 100h		

ENSEIGNANT(E) RESPONSABLE

FARINAS Jérôme

Email : jerome.farinas@irit.fr

Téléphone : 0561558343

LECHANI-TAMINE Lynda

Email : Lynda.Tamine-Lechani@irit.fr

Téléphone : 0561533881

PINQUIER Julien

Email : pinquier@irit.fr

Téléphone : 7434

OBJECTIFS D'APPRENTISSAGE

Etre capable de mettre en oeuvre des compétences en analyse et exploitation de données audio, vidéo et textuelle à un cas d'étude concret

Être capable de s'organiser et de suivre une méthodologie pour le travail en groupe.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Cas d'étude relatif à l'analyse et l'exploitation de données audio, vidéo et textuelles

Méthodologie de développement, méthodes fonctionnelles (SADT...)

Mise en application par groupe de 3-4 personnes

- recherche bibliographique
- développement d'un produit basé sur les méthodologies de la recherche biblio
- présentation sous forme de poster et d'un rapport

PRÉ-REQUIS

Extraction d'information (S7), Méthode de catégorisation/classification (K-means, KNN, SVM, EM)

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Modèles statistiques pour l'accès à l'information textuelle, Eric Gaussier, François Yvon
2. Multimodal Processing and Interaction. Audio, Video, Text, Springer, Petros Maragos
3. Structured Analysis Design Technic, Michel Galinier

MOTS-CLÉS

modélisation automatique, texte, audio, video, indexation par le contenu, méthodologie de développement, recherche bibliographique

UE	ANGLAIS	3 ECTS	1^{er} semestre
EIMAB3VM	TD : 24h		

ENSEIGNANT(E) RESPONSABLE

CHAPLIER Claire

Email : claire.chaplier@univ-tlse3.fr

OBJECTIFS D'APPRENTISSAGE

Niveau C1 du CECRL (Cadre Européen de Certification en Langues)

Développer les compétences indispensables aux étudiant/es en vue de leur intégration dans la vie professionnelle.

Perfectionner les outils de communication permettant de s'exprimer dans le contexte international d'aujourd'hui et acquérir l'autonomie linguistique nécessaire à cette intégration

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Contenu linguistique de la discipline :

Enseignement axé sur le travail de l'expression orale

Documents du domaine de spécialité pouvant faire l'objet de collaboration entre enseignants de science et enseignants de langue

Nécessité d'un parcours individualisé répondant aux attentes de chaque étudiant.

Compétences

CO - EE - EO - EE

- Savoir communiquer en anglais scientifique

- Savoir repérer les éléments constitutifs d'une communication écrite ou orale dans le domaine de spécialité

- Savoir prendre la parole en public (conférence ou réunion) dans le cadre d'un colloque, projet de recherche, projet professionnel

PRÉ-REQUIS

N/A

RÉFÉRENCES BIBLIOGRAPHIQUES

N/A

MOTS-CLÉS

Projet - Repérer - Rédaction anglais scientifique - style - registre - critique - professionnel - commenter

UE	NOSQL ET EXPERIENCES INDUSTRIELLES EN SCIENCES DES DONNEES	3 ECTS	2nd semestre
EIMAB4AM	Cours : 30h		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

HUBERT Gilles

Email : hubert@irit.fr

PITARCH Yoann

Email : yoann.pitarch@irit.fr

Téléphone : 05 61 55 6323

OBJECTIFS D'APPRENTISSAGE

1. NoSQL

Connaître les principes des bases de données NoSQL et apprendre à les manipuler.

2. Conférences thématiques

L'objectif des conférences thématiques est double :

- d'une part, d'offrir des compléments de formation pointus sur des sujets techniques précis de la formation,
- d'autre part, de présenter un état du marché du travail dans le secteur décisionnel et des data sciences.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

1. NoSQL

- Panorama des différents types de bases NoSQL
- Manipulation d'une base de données NoSQL

2. Conférences thématiques

Présentation par des industriels du secteur sur le marché du travail, des projets Big Data ou en Science des données.

PRÉ-REQUIS

Bases de données relationnelles, algorithmique et programmation, toutes les compétences du M1

RÉFÉRENCES BIBLIOGRAPHIQUES

Les bases de données NoSQL : Comprendre et mettre en oeuvre - Rudi Bruchez, Eyrolles

MongoDB - The Definitive Guide, Kristina Chodorow, O'Reilly

MOTS-CLÉS

NoSQL, MongoDB, HBase, Neo4j, Big data, data science, marché du travail

UE	PROJET / CHALLENGE OPEN-DATA OU BIG-DATA	6 ECTS	2nd semestre
EIMAB4BM	Projet : 150h		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

PINEL-SAUVAGNAT Karen

Email : Karen.Sauvagnat@irit.fr

Téléphone : 05 61 55 63 22

OBJECTIFS D'APPRENTISSAGE

Je sais mettre en œuvre les outils et concepts vus dans la formation sur un problème réel.

Je sais travailler en équipe et gérer un projet

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Projet de synthèse mobilisant l'ensemble des compétences de la formation.

PRÉ-REQUIS

Bases de données, programmation, statistiques...

RÉFÉRENCES BIBLIOGRAPHIQUES

<http://vizurbi.com/#{}>

<https://cmisid.univ-tlse3.fr/projet-taxisid>

MOTS-CLÉS

Projet de synthèse

UE	STAGE	21 ECTS	2nd semestre
EIMAB4CM	Stage : 5 mois minimum		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

HUBERT Gilles

Email : hubert@irit.fr

OBJECTIFS D'APPRENTISSAGE

Je sais mettre en oeuvre les connaissances et compétences acquises au cours de la formation pour répondre à différentes missions en science des données, dans un contexte applicatif industriel ou recherche.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Missions dans un contexte applicatif industriel ou recherche.

PRÉ-REQUIS

Ensemble de compétences mathématiques et informatiques dépendant du contexte applicatif.

MOTS-CLÉS

Mission, cadre applicatif, industrie, recherche

UE	STAGE	21 ECTS	2nd semestre
EIMAB4DM	Stage : 5 mois minimum		

ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : cecile.chouquet@math.univ-toulouse.fr

Téléphone : 05.61.55.69.84

HUBERT Gilles

Email : hubert@irit.fr

OBJECTIFS D'APPRENTISSAGE

Je sais mettre en oeuvre les connaissances et compétences acquises au cours de la formation pour répondre à différentes missions dans un contexte applicatif industriel ou recherche.

DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Missions dans un contexte applicatif industriel ou recherche.

PRÉ-REQUIS

Ensemble de compétences mathématiques et informatiques dépendant du contexte applicatif.

MOTS-CLÉS

Mission, cadre applicatif, industrie, recherche

GLOSSAIRE

TERMES GÉNÉRAUX

DÉPARTEMENT

Les départements d'enseignement sont des structures d'animation pédagogique internes aux composantes (ou facultés) qui regroupent les enseignants intervenant dans une ou plusieurs mentions

UE : UNITÉ D'ENSEIGNEMENT

Unité d'Enseignement. Un semestre est découpé en unités d'enseignement qui peuvent être obligatoire, optionnelle (choix à faire) ou facultative (UE en plus). Une UE représente un ensemble cohérent d'enseignements auquel est associé des ECTS.

ECTS : EUROPEAN CREDITS TRANSFER SYSTEM

Les ECTS sont destinés à constituer l'unité de mesure commune des formations universitaires de Licence et de Master dans l'espace européen depuis sa création en 1989. Chaque UE obtenue est ainsi affectée d'un certain nombre d'ECTS (en général 30 par semestre d'enseignement). Le nombre d'ECTS est fonction de la charge globale de travail (CM, TD, TP, etc.) y compris le travail personnel. Le système des ECTS vise à faciliter la mobilité et la reconnaissance des diplômes en Europe.

TERMES ASSOCIÉS AUX DIPLOMES

Les diplômes sont déclinés en domaines, mentions et parcours.

DOMAINE

Le domaine correspond à un ensemble de formations relevant d'un champ disciplinaire ou professionnel commun. La plupart de nos formations relèvent du domaine Sciences, Technologies, Santé.

MENTION

La mention correspond à un champ disciplinaire. Elle comprend, en général, plusieurs parcours.

PARCOURS

Le parcours constitue une spécialisation particulière d'un champ disciplinaire choisie par l'étudiant au cours de son cursus.

TERMES ASSOCIÉS AUX ENSEIGNEMENTS

CM : COURS MAGISTRAL(AUX)

Cours dispensé en général devant un grand nombre d'étudiants (par exemple, une promotion entière), dans de grandes salles ou des amphis. Au-delà de l'importance du nombre d'étudiants, ce qui caractérise le cours magistral, est qu'il est le fait d'un enseignant qui en définit lui-même les structures et les modalités. Même si ses contenus font l'objet de concertations entre l'enseignant, l'équipe pédagogique, chaque cours magistral porte la marque de l'enseignant qui le dispense.

TD : TRAVAUX DIRIGÉS

Ce sont des séances de travail en groupes restreints (de 25 à 40 étudiants selon les composantes), animés par des enseignants. Ils illustrent les cours magistraux et permettent d'approfondir les éléments apportés par ces derniers.

TP : TRAVAUX PRATIQUES

Méthode d'enseignement permettant de mettre en pratique les connaissances théoriques acquises durant les CM et les TD. Généralement, cette mise en pratique se réalise au travers d'expérimentations. En règle générale, les groupes de TP sont constitués des 16 à 20 étudiants. Certains travaux pratiques peuvent être partiellement encadrés voire pas du tout. A contrario, certains TP, du fait de leur dangerosité, sont très encadrés (jusqu'à 1 enseignant pour quatre étudiants).

PROJET OU BUREAU D'ÉTUDE

Le projet est une mise en pratique en autonomie ou en semi-autonomie des connaissances acquises. Il permet de vérifier l'acquisition des compétences.

TERRAIN

Le terrain est une mise en pratique encadrée des connaissances acquises en dehors de l'université.

STAGE

Le stage est une mise en pratique encadrée des connaissances acquises dans une entreprise ou un laboratoire de recherche. Il fait l'objet d'une législation très précise impliquant, en particulier, la nécessité d'une convention pour chaque stagiaire entre la structure d'accueil et l'université.

