

Licence Informatique "Big Data"

Programme Détaillé



Licence 1 – Semestre 1

Anglais

Dans ce cours nous travaillerons sur les principales caractéristiques de communication en Anglais, à l'oral et à l'écrit. L'objectif est de permettre aux étudiants d'avoir l'équivalent du niveau "Beginning Literacy ESL (English as a Second Language)".

Expression et Communication

Ce cours est conçu pour permettre aux étudiants d'avoir de bonnes bases en communication orale pour pouvoir assurer des présentations individuelles ou collectives dans un cadre professionnel. L'accent sera mis sur la révision des règles de grammaire de base, les techniques de présentation synthétique de projet et les techniques de compte rendu.

Système Unix et installation Linux

Ce cours permet de comprendre des bases du fonctionnement du système d'exploitation Unix et ses outils à travers une distribution Linux. Après la présentation de Linux, la gestion des fichiers et édition ainsi que la programmation en Shell seront abordées.

Bases du Web (Html, CSS, JS, Hébergement, DNS)

Dans ce cours les étudiants exploreront les différentes fonctionnalités des langages de base du web : HTML, du CSS et javascript. L'objectif est qu'ils puissent comprendre les différentes notions et savoir quand utiliser un langage plutôt qu'un autre. Ainsi, ils pourront être totalement autonomes et seront capables de mettre en place des pages web dynamiques.

Licence Informatique "Big Data"

Programme Détaillé



Licence 1 - Semestre 1

Langage Python

Python est le langage de programmation le plus utilisé dans le domaine du Big Data et de la Data Science. Dans ce cours l'objectif sera de donner aux étudiants une connaissance active de l'algorithme de base et de la programmation structurée; leur permettre de comprendre les concepts de base du langage et de l'utiliser pour résoudre les problèmes posés. L'accent sera mis sur le programme, les types de données et opérateurs et les manipulations complexes des listes et dictionnaires.

Langage R

Apprentissage du logiciel statistique R et au langage informatique du même nom afin de l'utiliser de manière autonome. Apprentissage des manipulations élémentaires pour pouvoir modifier et organiser des données brutes dans le but de les utiliser lors d'analyses statistiques. Il s'agira également d'apprendre à définir ses propres fonctions, de les appliquer à des variables et de s'initier aux boucles dans R. L'apprentissage de la création graphique avec R sera ensuite effectué.

Algorithmique

Dans le cadre de ce cours est fourni le cadre méthodologique permettant de comprendre les approches algorithmiques, la complexité, le concept de variables, de structures de contrôle, de structures de données afin de concevoir et écrire des algorithmes.

Licence Informatique "Big Data"

Programme Détaillé



Licence 1 – Semestre 1

Bases de données SQL

Introduction aux concepts de Base de données et de SGBD (Système de Gestion de Bases de Données) de type « Relationnel » suivi de la présentation du modèle Entité-Relation et du modèle relationnel CODD en mettant l'accent sur la conception d'un bon schéma de base de données (dépendances fonctionnelles, théorie de la normalisation,...) Le langage SQL est aussi étudié.

Stockage de données dans le cloud

Ce cours abordera les principaux problèmes posés et les méthodes proposées dans la conception des systèmes cloud ainsi que les concepts et les termes à connaître pour exploiter Google Cloud Platform (GCP). Les nombreux services informatiques et services de stockage disponibles sur cette dernière seront présentés: Google App Engine, Google Compute Engine, Google Kubernetes Engine, Google Cloud Storage, Google Cloud SQL et BigQuery.

Licence Informatique "Big Data"

Programme Détaillé



Licence 1 – Semestre 2

Anglais

Le cours d'anglais de ce semestre est l'équivalent du niveau "Low Beginning ESL". L'accent sera mis sur les expressions fréquentes dans la Tech, les tâches de routine (à travers le prisme de Python et des modules les plus utilisés en Data Science) et l'expression des besoins de base.

Expression et communication

Ce module insistera principalement sur les techniques de communication orale. L'objectif sera de permettre aux étudiants maîtriser l'énoncé d'un problème et l'argumentation à travers des exercices pratiques.

Outils statistiques pour la Data Science

Familiarisation avec la méthode statistique en général, en vue de décrire, de résumer et d'analyser une population ou un ensemble de données. L'objectif sera de maîtriser les concepts de base (terminologie) de la statistique descriptive, ainsi que les notions essentielles sur les distributions statistiques (à un et deux caractères), les indices et les séries chronologiques. L'étudiant devra ainsi se familiariser avec les méthodes de traitement et d'analyse de l'information chiffrée, étudier les phénomènes, les expliquer, les prévoir pour pouvoir prendre les décisions les plus appropriées, ce qui constitue un atout pour l'insertion professionnelle.

Langage Python

Ce cours permet de comprendre et de concevoir un algorithme mais aussi d'évaluer son efficacité. Les étudiants pourront écrire un programme et manipuler les objets et les structures de données.

Licence Informatique "Big Data"

Programme Détaillé



Licence 1 – Semestre 2

DevOps

Ce cours abordera la méthodologie DevOps, les outils DevOps et l'architecture système. L'objectif est de familiariser les étudiants avec l'approche DevOps et leur apprendre les outils nécessaires à l'automatisation des installations et du cycle de vie d'une application, la mise en place d'une chaîne d'intégration continue et de containerisation.

Initiation IoT – Arduino

L'internet des objets constitue l'ensemble des dispositifs électroniques connectés à Internet. Ces dispositifs permettent de collecter des données de toutes sortes permettant de mesurer l'environnement et de gérer des interactions avec un utilisateur. Lors de ce cours les étudiants seront initiés à l'utilisation du microcontrôleur Arduino via la programmation de celui-ci. Ils seront amenés à utiliser les capteurs avec Arduino, à gérer l'interface d'entrée (boutons, clavier) et les affichages en sortie.

Outils IA en ligne

Ce module abordera les tableaux Software, le DialogFlow et le DataWrapper. L'objectif est la familiarisation avec les capacités du Big Data et de l'IA via l'utilisation d'outils en ligne ne requérant pas de connaissances en programmation ; mais également de permettre aux étudiants d'appréhender les différentes applications possibles du Big Data de manière opérationnelle.

Projet informatique

L'objectif de ce projet est double : d'une part, il permet de renforcer les notions vues en cours ; d'autre part, il constitue un premier contact avec le travail en équipe. Les projets sont réalisés par groupes de 2 ou 3 étudiants. Ce module vise à former à la résolution d'un problème scientifique et technique concret nécessitant l'élaboration d'une application informatique. Il vise également à apprendre à travailler en équipe et à savoir présenter son travail de manière claire et concise, sous la forme d'un rapport écrit et d'une soutenance orale à l'aide d'un diaporama.

Licence Informatique "Big Data"

Programme Détaillé



Licence 2 – Semestre 3

Anglais

Le cours d'anglais de ce semestre est l'équivalent du niveau "High Beginning ESL". L'accent sera mis sur l'organisation de l'information et les questions simples liées aux besoins de "survie" et aux échanges sociaux simples; l'écoute active et la compréhension de conversations simples; la lecture incluant la prévisualisation, le visionnage et la prédiction; la production de textes écrits simples.

UX Design

Ce cours permet de connaître les outils pertinents pour les concepteurs UX et la différence entre UI et UX. La conception UX est une technique qui crée des produits, le plus souvent des applications et des sites Web, faciles à utiliser et répondant aux exigences de l'utilisateur final. Il s'agit de comprendre ce dont l'utilisateur cible a besoin et comment il obtient ce qu'il veut. Les étudiants seront capables de créer des profils d'utilisateurs et des personnages et d'interagir avec différents types d'utilisateurs.

Outils mathématiques pour la Data Science

Formulation mathématique de problèmes concrets simples. Apprentissage des notions de base sur les relations, l'algèbre de Boole et les fonctions booléennes. Calculs simples sur les dénombrements et les probabilités combinatoires, la récurrence. Etude des rudiments d'arithmétique.

Langage Python

Ce cours présentera l'environnement de programmation python approfondie. Les étudiants se familiariseront avec les techniques de manipulation et de nettoyage des données à l'aide de la bibliothèque de science des données python pandas. Ils aborderont aussi l'abstraction de la série et du DataFrame en tant que structures de données centrales pour l'analyse des données, ainsi que des didacticiels sur la façon d'utiliser des fonctions telles que groupby, merge et le pivotage efficace des tableaux. À la fin du module, les étudiants seront capables de prendre des données tabulaires, de les nettoyer, de les manipuler et d'effectuer des analyses statistiques inférentielles de base.

Licence Informatique "Big Data"

Programme Détaillé



Licence 2 – Semestre 3

Bases de données no-SQL

Ce module permet de connaître l'univers et les principes des bases de données NoSQL. L'étudiant apprendra la gestion d'une énorme base de données et son interrogation de manière efficace. Le focus sera mis sur les solutions MongoDB et ElasticSearch. L'objectif est de permettre à l'étudiant d'être capable de stocker et de réaliser des requêtes sur les données en grande dimension tout en assurant le passage à l'échelle.

Data collection

Ce cours donne un aperçu des produits de données existants et une bonne compréhension du paysage de la collecte de données. Il fournira aussi un cadre général qui permettra non seulement de comprendre chaque étape requise pour une collecte et une analyse de données réussies, mais aussi d'aider à identifier les erreurs associées à différentes sources de données.

Data Engineering 1

Dans ce module, l'étudiant sera exposé au monde de l'ingénierie des données. Il aura un aperçu des divers outils que les ingénieurs de données utilisent et approfondira sa compréhension de la façon dont la technologie cloud joue un rôle dans l'ingénierie des données. Il travaillera sur l'informatique parallèle qui est la pierre angulaire des outils de l'ingénieur de donnée, ainsi que sur la planification des travaux de traitement de données. Il découvrira ensuite comment extraire, transformer et charger des données avec la méthode ETL.

Infosec

Dans ce cours, l'étudiant aborde la question de la sécurité de l'information de manière théorique et technique. Des séances de travaux pratiques permettent d'assimiler les outils nécessaires aussi bien en audit de sécurité qu'en cryptographie. À la fin du cours, l'étudiant est capable d'identifier les vulnérabilités d'un système, de proposer des solutions et de former ses collaborateurs en transmettant les bonnes pratiques à avoir.

Licence Informatique "Big Data"

Programme Détaillé



Licence 2 – Semestre 4

Anglais

Le cours d'anglais du quatrième semestre est l'équivalent du niveau "Low Intermediate ESL". L'accent sera mis sur la lecture d'un journal technique et la conversation orale.

Bases de la comptabilité

Ce cours constitue une introduction pour les débutants et une remise à niveau pour ceux qui ont déjà une formation comptable. Sachant que notre formation est grandement axée vers l'entrepreneuriat, ce cours sert à acquérir et consolider les connaissances fondamentales qui sont essentielles à la gestion d'une petite entreprise.

Outils statistiques pour la Data Science

Dans ce module il s'agira de présenter les techniques modernes de l'analyse de grands ensembles de données et développer les outils de base de l'analyse de données et du "Data mining". En plus de rappels sur l'algèbre et la géométrie, l'étudiant apprendra l'analyse de données et les méthodes de classification.

Outils mathématiques pour la Data Science

Ce cours vise à donner à l'étudiant une palette d'outil mathématiques utiles à sa discipline. Les polynômes, les espaces vectoriels, la diagonalisation de matrices complexes et les outils d'analyse seront étudiés. Tout en assurant la rigueur propre aux mathématiques de chacun des sujets présentés, nous souhaitons maintenir une interaction constante entre les problèmes concrets et leur formalisation.

Licence Informatique "Big Data"

Programme Détaillé



Licence 2 - Semestre 4

Initiation au Machine Learning

Après la collecte, le nettoyage et l'exploration des données, on passe à la partie de modélisation statistique.

C'est le processus que nous allons explorer dans ce module consacré au Machine Learning, qui est un ensemble de techniques puissantes permettant de créer des modèles prédictifs à partir de données, sans avoir été explicitement programmées.

Ce cours permettra de décrire le travail général du data scientist, de définir le machine learning, son champs d'application et ses limites, d'identifier les différents types de modélisation statistique (classification, régression, apprentissage supervisé, apprentissage non-supervisé).

Containers et virtualisation

La virtualisation est une technique qui a révolutionné les technologies de l'information. Ce module permet aux étudiants de découvrir les concepts et impacts de la virtualisation,

les différentes technologies possibles sur le marché ainsi que leurs avantages/inconvénients, et la manière de mener un projet de virtualisation.

Après l'introduction à la virtualisation et aux containers, les étudiants apprendront la virtualisation du stockage et les solutions d'administration et de supervision.

Architecture Big Data

Ce cours donne un aperçu des offres technologiques essentielles et de la valeur résultante des composants Big Data.

Nous ferons une introduction à l'architecture lambda et étudierons les méthodes de Batch Processing et de Speed Processing.

Licence Informatique "Big Data"

Programme Détaillé



Licence 2 - Semestre 4

Visualisation des données

L'objectif de ce cours centré sur les techniques de visualisation de grands volumes de données, est de présenter des techniques de visualisation d'information et de les appliquer aux problèmes de représentation et d'interaction avec de grands volumes de données. À la fin de ce cours les étudiants seront capables de décrire les principales techniques de visualisation et de les développer à l'aide d'un framework pour la programmation de techniques de visualisation sur le Web.

Projet Informatique

Ce module vise à former à la résolution d'un problème scientifique et technique concret nécessitant l'élaboration d'une application informatique. Il vise également à apprendre à travailler en équipe et à savoir présenter son travail de manière claire et concise, sous la forme d'un rapport écrit et d'une soutenance orale à l'aide d'un diaporama. Les projets proposés sont choisis dans des domaines variés.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 5

Anglais

L'objectif de ce cours est la réussite du TOEIC ("Test of English for International Communication"): un test de compétence pour les anglophones non natifs. Le TOEIC mesure la capacité de communication en anglais dans les affaires, le commerce, et l'industrie. Ce semestre sera l'équivalent du niveau "High Intermediate ESL".

Management des données et GDPR

Le respect de la GDPR (General Data Protection Regulation), la plus grande loi sur la protection de la vie privée au monde, est devenu un critère obligatoire pour toutes les entreprises qui traitent les données des citoyens européens à l'intérieur et à l'extérieur de l'UE. Notions sur le Digital Markets Act (DMA). Dans ce module il y aura une vue d'ensemble de la loi, de ses principes, droits et responsabilités.

Start-Up / Jeu Entreprise

La création d'une start-up ne consiste pas simplement à élaborer un plan d'exécution pour un modèle d'entreprise qui, selon l'entrepreneur, fonctionnera, mais plutôt à rechercher le modèle d'entreprise réel lui-même. L'idée principale de ce cours est d'apprendre à développer et à tester rapidement des idées en recueillant des quantités massives de réactions des clients et du marché.

Chatbot

Le cours débute par l'étude de l'architecture d'un système complexe de traitement automatique du langage (reconnaissance, analyse, génération). Il se poursuit par l'étude de théories linguistiques et formalismes informatiques centraux dans le domaine du NLP. Une attention particulière est accordée la présentation et l'analyse d'applications réelles.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 5

DevOps

Dans ce cours, l'étudiant approfondit les connaissances du précédent module. Il acquiert les outils nécessaires à l'optimisation de la livraison d'un produit. À la fin du cours, il doit pouvoir mettre en place une démarche DevOps en entreprise, optimiser les performances d'un système informatique et assurer la disponibilité des infrastructures.

Initiation à la programmation Parallèle

Ce module est une initiation aux principaux paradigmes de la programmation parallèle. À la fin du cours, l'étudiant devrait être capable d'expliquer et d'utiliser les principales stratégies de programmation parallèle ; lire des programmes parallèles écrits dans divers langages ; développer des programmes (en Threaded-C) fonctionnant sur une véritable machine parallèle et expliquer les forces et faiblesses des différents paradigmes.

Traitement du signal

Avec une approche basée sur la programmation, ce cours est conçu pour donner une base solide dans les aspects les plus utiles du traitement numérique du signal (DSP) d'une manière engageante et facile à suivre.

L'objectif de ce cours est de présenter des techniques pratiques tout en évitant les obstacles des théories mathématiques abstraites. Les techniques DSP sont expliquées dans un langage simple.

Data Engineering

Ce cours vous montrera les bases de l'apprentissage automatique pour les ingénieurs de données.

Ce n'est pas un cours général ou une introduction à l'apprentissage automatique. Il s'agit d'un cours très ciblé pour apprendre les concepts de l'ingénierie des données.

La construction de modèles en Python, les méthodes Scala et Spark et la programmation Kubernetes Scala seront abordées.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 5

Gestion d'un Data Warehouse

a Ce cours s'inscrit dans le domaine de " l'informatique décisionnelles (ID) " ou "Business Intelligence (BI)" en Anglais. Il traite des entrepôts de données - ED (Data Warehouse). Après avoir défini ce que recouvrait la DI/BI, est introduit le concept d'ED et l'analyse en ligne OLAP. On s'intéresse ensuite à la conception d'un ED tant à un niveau conceptuel qu'à un niveau logique. Les grandes approches d'implantation des ED sont présentées: l'approche ROLAP, l'approche MOLAP, et l'approche hybride HOLAP. Des langages informatiques permettant de faire l'analyse OLAP d'ED sont introduits.

Initiation au Machine Learning (NLP)

Ce cours présente le traitement du langage naturel grâce à l'utilisation de python et du kit d'outils en langage naturel. Grâce à une approche pratique, il permet d'acquérir une expérience pratique de l'utilisation et de l'analyse de texte. il aborde une variété de façons de représenter les langues humaines en tant que systèmes informatiques et comment exploiter ces représentations pour écrire des programmes qui font des choses utiles avec des données de texte et de parole, comme la traduction, le résumé, l'extraction d'informations, la réponse aux questions, les interfaces naturelles avec les bases de données et les agents conversationnels.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 6

Anglais

Le cours d'anglais de ce semestre est l'équivalent du niveau "Advanced ESL". Il s'agira de s'initier au Génie en herbe; d'apprendre à résumer les informations provenant de différentes sources orales et écrites, en reconstituant les arguments et les compes dans une présentation cohérente; et de s'exercer sur la résolution des conflits.

L'art du Pitch

Le pitch est primordial et savoir comment lancer une start-up est essentiel. Peu importe que l'idée de produit soit innovante, bien pensée ou potentiellement rentable: si l'on ne parvient pas à attirer des investisseurs dans sa start-up, l'entreprise aura de sérieuses difficultés à se développer. Ce cours permettra à l'étudiant de découvrir ce que les investisseurs recherchent; d'apprendre à se présenter avec succès aux investisseurs, de savoir prendre un raccourci avec les experts en la matière et de maîtriser le pitch de start-up.

Gestion de projet Agile

La méthodologie agile est un type de processus de gestion de projet, principalement utilisé pour le développement de logiciels, où les demandes et les solutions évoluent grâce à l'effort de collaboration des équipes auto-organisées et interfonctionnelles et de leurs clients. Ce cours permettra à l'étudiant de maîtriser les valeurs et les principes du "Manifeste pour le développement logiciel agile"; de maîtriser la chaîne des six principaux "produits livrables" pour suivre les progrès et créer le produit; de se familiariser avec le cadre agile, Scrum et d'apprendre la correcte implémentation de la méthodologie agile.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 6

Visualisation des données

Ce cours développe les concepts généraux de l'exploration de données ainsi que les méthodologies et applications de base. Ensuite, il se focalise sur le sous-champ de l'exploration de données : les modèles. Ce cours offre à l'étudiant la possibilité d'acquérir des compétences sur les méthodes évolutives de découverte de modèles sur des données transactionnelles massives, discuter des mesures d'évaluation et étudier des méthodes pour explorer divers types de modèles (séquentiels et modèles de sous-graphique).

Outils d'aide à la Décision

Le cours permettra aux étudiants de comprendre la Business Intelligence (BI) qui est une catégorie d'applications et de technologies pour la collecte, le stockage, l'analyse, le partage et l'accès aux données pour aider les utilisateurs d'entreprise à prendre de meilleures décisions de gestion. Ils apprendront les principes et les meilleures pratiques pour l'utilisation des données afin de soutenir la prise de décision basée sur les faits. L'accent sera mis sur les applications en marketing, où la BI aide, par exemple, à analyser les retours de campagne, les rendements promotionnels ou le suivi du marketing sur les réseaux sociaux ; dans les ventes, où la BI aide à effectuer l'analyse des ventes ; et dans les domaines d'application tels que la gestion de la relation client et le commerce électronique.

Projet Informatique

Ce module vise à former à la résolution d'un problème scientifique et technique concret nécessitant l'élaboration d'une application informatique. Il vise également à apprendre à travailler en équipe et à savoir présenter son travail de manière claire et concise, sous la forme d'un rapport écrit et d'une soutenance orale à l'aide d'un diaporama.

Les projets proposés sont choisis dans des domaines variés.

Licence Informatique "Big Data"

Programme Détaillé



Licence 3 – Semestre 6

Stage

Le stage a pour but de fournir aux étudiants une formation complémentaire à la formation académique en confrontant le stagiaire aux exigences d'une production avec l'encadrement d'un milieu professionnel établi. Le stage constitue une étape transitoire entre le milieu scolaire et le milieu professionnel. L'application des apprentissages scolaires dans le milieu de travail permet le plus souvent de se familiariser avec le poids de la hiérarchie dans les organisations; de sentir aussi les effets d'une stricte division du travail ; de découvrir à la fois les réseaux complexes de ressources disponibles et les réglementations à observer. C'est une méthode active de perfectionnement de l'apprentissage pratique dans un contexte non simulé. Les stages sont proposés par les partenaires professionnels du Dakar Institute of Technology. Ces partenaires sont des entreprises, des ONG et des structures étatiques. L'étudiant se trouve ainsi dans un milieu professionnel reconnu.