Code UE	Durée (H)	Responsable(s)	Entreprise Resp.
DE01	10	Babiga Birregah	Université de Technologie de Troyes
DE02	30	Mohammed Lahlou/ Lyes Khoukhi	Africa Data Lab/ Université de Technologie de Troyes
DE03	40	Mohammed Lahlou	Africa Data Lab
DE04	20	Andrea Duhamel	Université de Technologie de Troyes
DE05	30	Ghita Ammor	Africa Data Lab
DE06	30	Ghita Ammor	Africa Data Lab
DE07	20	Samiha Ayed	Université de Technologie de Troyes
	180		

Intitulé
Introduction au Big Data
Base de données avancées (NoSQL, graphes, etc.)
Ecosystème Hadoop
Intelligence artificielle et Machine Learning
Big Data : applications en batch
Big Data: Applications en streaming
Conception et gestion d’un projet Big Data

Contenu
1.1 : Présentation de l'écosystème Big Data
1.2 : Méthodologies et process de mise en œuvre
1.3 : Enjeux et gouvernance de la donnée en entreprise
2.1 : Comparatif des différents types de Bases de Données (structurées, semi-structurées, et non structurées)
2.2 : Architecture des différentes bases NoSQL
2.3 : Les bases NoSQL clé/valeur, documents, graphes
3.1 : Présentation de l'écosystème Hadoop
3.2 : Architecture de projets Big Data
3.3 : Conception de projets Big Data hautement parallélisés
3.4 : Manipulation des données avec Hive, Flume
4.1 : Introduction au Data Mining
4.2 : Mise en forme des données : Feature Engineering
43 : Présentation des différents algorithmes de machine learning
4.4 : Industrialisation des modèles et passage à l'échelle
5.1 : Développement et optimisation d'applications avec Apache Spark
5.2 : Création d'applications distribuées avec Java en se basant sur le paradigme MapReduce.
5.3 : Automatisation et orchestration des traitements avec Oozie
6.1 : Déploiement et gestion d'un cluster pour traiter les données en streaming
6.2 : Gestion des applications en temps réel avec Kafka
6.3 : Développement d'applications avec Apache Spark Streaming
7.1 : Comprendre les spécificités technologiques, organisationnelles et juridiques d'un projet Big Data
7.2 : Mise en place et pilotage d'une équipe projet Big Data
7.3 : Problématiques de protection et de sécurité des projets Big Data