le cnam
Nouvelle-Aquitaine

Fiche récapitulative

STA211 | Entreposage et fouille de données

78
Total d'heures d'enseignement

9
Crédits ECTS

02/02/2026
Début des cours prévu

Programme

Modèles prévisionnels et systèmes de gestion de l'entreprise
- structures spécifiques des bases de données de Data warehouse (star schema)
- OLAP
Méthodologies générales
- Méthodologies de Data Mining
Pré-traitement des données
- Analyses de la qualité des données,
- Techniques d'appréhension des valeurs manquantes ou aberrantes
- Techniques de construction de bases de travail (agrégations, etc. . .)
Données et techniques de fouille
Méthodes non supervisées :
- Cartes de Kohonen, Règles d'association
Méthodes supervisées :
- Rappels de théorie de l'apprentissage
- Arbres de décision, forêts aléatoires, SVM, Réseaux de neurones, deep learning
- Méta-algorithmes :
- boosting, bagging
Fouille dans de nouveaux types de données et méthodes associées :
- Données textuelles - Données multivues - Images et Multimedia
Outils :
- Environnements freeware : R, Python
- Outils spécifiques : SAS-EM, SPAD
- Data Mining et bases de données : OLAP Business Object

Objectifs : aptitudes et compétences

Compétences :

Compétences développées

1. Modèles prévisionnels et systèmes de gestion de l'entreprise

- Compétences : Conception et gestion de Data Warehouses en utilisant des structures de base de données adaptées comme le schéma en étoile (star schema). Compréhension des principes de l'analyse multidimensionnelle avec OLAP.
- Exemple d'application : Construction d?un entrepôt de données pour une entreprise, permettant d'extraire des indicateurs clés pour l?aide à la décision à l?aide de tableaux croisés OLAP.

2. Méthodologies générales de Data Mining

- Compétences : Maîtriser les méthodologies de Data Mining pour l'exploration, la classification et la modélisation des données.
- Exemple d'application : Utilisation de techniques de fouille pour détecter des patterns cachés dans des données clients et améliorer les stratégies marketing.

3. Pré-traitement des données

- Compétences : Améliorer la qualité des données grâce à des techniques de pré-traitement comme le nettoyage des données, la gestion des valeurs manquantes ou aberrantes, et la construction de bases de travail par agrégation.

- Exemple d'application : Nettoyage d?un ensemble de données client en supprimant les valeurs manquantes, puis agrégation des données pour réaliser des analyses.

4. Données et techniques de fouille

Méthodes non supervisées :

- Compétences : Utiliser des techniques comme les cartes de Kohonen et les règles d'association pour découvrir des structures dans les données sans avoir besoin de labels.
- Exemple d'application : Application des cartes de Kohonen pour segmenter automatiquement des groupes de clients en fonction de leurs comportements d?achat.

Méthodes supervisées :

- Compétences : Maîtriser des algorithmes d'apprentissage supervisé, tels que les arbres de décision, les forêts aléatoires, les réseaux de neurones et le deep learning, pour la classification et la prédiction.
- Exemple d'application : Utilisation de forêts aléatoires pour prédire la probabilité qu'un client souscrive à un service basé sur son historique d?achat.

5. Méta-algorithmes : Boosting, Bagging

- Compétences : Maîtriser les techniques de méta-algorithmes comme le boosting et le bagging pour améliorer la performance des modèles de machine learning.
- Exemple d'application : Utilisation de boosting pour combiner plusieurs modèles de faible performance en un modèle global plus performant dans une tâche de classification.

6. Fouille de nouveaux types de données et méthodes associées

- Compétences : Analyser et extraire de la valeur de nouveaux types de données, comme les données textuelles, multivues, et images/multimedia.
- Exemple d'application : Application de méthodes de fouille pour analyser des avis clients sous forme de texte ou classer des images dans une base de données multimédia.

7. Outils et environnements logiciels

- Compétences : Utiliser des environnements de programmation open-source comme R et Python pour la fouille de données et l? analyse statistique. Maîtriser les outils spécifiques comme SAS-EM, SPAD pour des analyses avancées de données, ainsi que l'intégration avec des bases de données via OLAP Business Object.
- Exemple d'application : Utilisation de R pour développer un modèle prédictif basé sur les données historiques de ventes, ou de SAS-EM pour appliquer des analyses statistiques avancées sur des données d'entreprise.

Aptitudes développées

1. Autonomie et prise d'initiative

- Aptitudes : Capacité à identifier et à implémenter les algorithmes et outils appropriés pour répondre à un besoin d'analyse de données spécifique.

2. Rigueur dans la gestion des données

- Aptitudes : S'assurer que les processus de pré-traitement des données sont effectués de manière rigoureuse pour garantir la qualité des données utilisées dans les modèles.

3. Capacité d'analyse et d'interprétation

- Aptitudes : Analyser de manière critique les résultats issus des modèles de fouille de données et les interpréter pour générer des informations exploitables pour l?entreprise.

4. Adaptabilité aux nouvelles technologies

- Aptitudes : Capacité à s'adapter rapidement à de nouveaux outils et environnements logiciels, ainsi qu'aux évolutions des algorithmes de Data Mining.

5. Gestion de la complexité

- Aptitudes : Capacité à gérer des ensembles de données volumineux et complexes, tout en appliquant des techniques adaptées de fouille et de pré-traitement.

6. Travail en équipe

- Aptitudes : Collaboration avec d'autres équipes (analystes de données, statisticiens, ingénieurs) pour intégrer les résultats d'analyses dans des processus métiers de l?entreprise.

7. Curiosité et veille technologique

- Aptitudes : Suivre les évolutions des méthodes de fouille de données et des technologies émergentes, telles que le deep learning et l'analyse des données non structurées.

Prérequis

être admis en M2 du master STIC, mention statistique ou être agréé.
La demande d'agrément est à faire uniquement pour les auditeurs qui souhaitent suivre STA211 en présentiel au semestre 1. Pour obtenir l'agrément, les auditeurs adresseront, à partir du 25 Août, par courrier électronique à l'enseignant responsable : ndeye.niang_keita@cnam.fr, un CV détaillé et une lettre de motivation indiquant les raisons de la demande et le projet pédagogique dans lequel elle s'inscrit . Une réponse sera donnée dans un délai d'une dizaine de jours.
Niveau requis : NFA008 (bases de données) et STA101 (analyse des données, méthodes descriptives). Ces prérequis sont indispensables pour obtenir l'agrément. Vérifier avant d'envoyer une demande d'agrément.

Délais d'accès

Le délai d'accès à la formation correspond à la durée entre votre inscription et la date du premier cours de votre formation.

- UE du 1er semestre et UE annuelle : inscription entre mai et octobre
- UE du 2e semestre : inscription de mai jusqu'à mi-mars

Exemple : Je m'inscris le 21 juin à FPG003 (Projet personnel et professionnel : auto-orientation pédagogique). Le premier cours a lieu le 21 octobre. Le délai d'accès est donc de 4 mois.

Planning

Légende:

Cours en présentiel

Cours 100% à distance

Mixte: cours en présentiel et à distance

Modalités	Lieux	Disponibilités	Prochaines sessions *	Tarif indicatif
	En ligne	Semestre 2	02/02/2026	De 0 à 1.560 €
	En ligne	Semestre 2	Prévue en 2026-2027	De 0 à 1.560 €
	En ligne	Semestre 2	Prévue en 2027-2028	De 0 à 1.560 €

*Selon les UEs, il est possible de s'inscrire après le début des cours. Votre demande sera étudiée pour finaliser votre inscription.

Modalités

Modalités pédagogiques :

Pédagogie qui combine apports académiques, études de cas basées sur des pratiques professionnelles et expérience des élèves. Équipe pédagogique constituée pour partie de professionnels. Un espace numérique de formation (ENF) est utilisé tout au long du cursus.

Modalités de validation :

Projet / mémoire

Pour valider cette UE, vous devez obtenir une note minimale de 10/20

Tarif

Mon employeur finance	1.560 €
Pôle Emploi finance	780 €
Je finance avec le co-financement Région	Salarié : 234 €
Je finance avec le co-financement Région	Demandeur d'emploi : 187,20 €

Plusieurs dispositifs de financement sont possibles en fonction de votre statut et peuvent financer jusqu'à 100% de votre formation.

Salarié : Faites financer votre formation par votre employeur

Demandeur d'emploi : Faites financer votre formation par Pôle emploi

Votre formation est éligible au CPF ? Financez-la avec votre CPF

Si aucun dispositif de financement ne peut être mobilisé, nous proposons à l'élève une prise en charge partielle de la Région Nouvelle-Aquitaine avec un reste à charge. Ce reste à charge correspond au tarif réduit et est à destination des salariés ou demandeurs d'emploi.

Pour plus de renseignements, consultez la page Financer mon projet formationopen_in_new ou contactez nos conseillers pour vous accompagner pas à pas dans vos démarches.

Cette unité d'enseignement n'est valorisable que dans cette certification.

Avis des auditeurs

Les dernières réponses à l'enquête d'appréciation de cet enseignement :

↓ Fiche synthétique au format PDF

Taux de réussite

Les dernières informations concernant le taux de réussite des unités dcenseignement composant les diplômes

↓ Taux de réussite