

KAÏNA-COM

CATALOGUE DE FORMATION

Les bases de la science des données (Data Science Essentials)

Comprendre les outils disponibles pour le "Data Scientist" en Python, les meilleures pratiques et les modèles de conception et "Machine Learning"



KBP001 – Data Science Essentials

Référence KBP001

Niveau

- Débutant
- Intermédiaire
- Expert

Nombre de Jours Programme de formation (160 H) :

- 40 x 4h par jour
 - Partie A - Python Développement (25 H)
 - Partie B – l’analyse des données avec python (Data Analysis with Python (25 H))
 - Partie C – Machine Learning (60 H)
 - Partie D – Exemple approfondi (Advanced Topics (50 H))

Lieu de la formation

- I: e-learning, Formation individuelle (Formation en Ligne)
- V: v-learning, classe virtuelle
- C: c-learning, cours présentiel

KAINA-COM
LE CARRÉ HAUSSMANN II,
6 Allée de la Connaissance
77127 Lieusaint – France

Prérequis La connaissance basique d'un langage objet est préférable.
Un niveau d'anglais business moyen est requis car la formation sera dispensée en anglais.

Public Manager de haut niveau, Manager avant-vente, Manager informatique, QA et support technique, programmeur et étudiant souhaitant en savoir plus sur la cybersécurité et l'IA.

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Objectifs

Part A - Python Développement

- Dans cette section, nous apprendrons le langage de développement python qui est actuellement un des langages les plus utilisés dans le monde. Nous étudierons en profondeur toutes les bases de ce langage, des sujets de syntaxe simples aux modules uniques pour ce langage. Les participants de ce module pourront, à l'issue de cette session, programmer et développer en langage Python. Ils auront ainsi le niveau requis pour l'industrie en général et pour le monde des traitements des données en particulier.

Part B – Analyses de données avec Python

- Dans cette section, nous nous concentrerons sur une variété de bibliothèques du langage Python les plus courantes utilisées pour le traitement et l'analyse de l'information : de la collecte de données, de l'analyse, de l'optimisation des modèles et de la présentation visuelle. Nous allons introduire des utilisations mathématiques et scientifiques des bibliothèques NumPy, scipy et Matplotlib

Part C - Machine Learning

- Dans ce module, nous étudierons en profondeur le sujet le plus pointu dans le monde de la science des données et de l'intelligence artificielle : l'apprentissage automatique. Ce qui permet aux ordinateurs de fonctionner « indépendamment » sans leur programmation. Dans ce module, nous analyserons toutes les différentes techniques, outils et algorithmes de cette zone importante

Part D – Exemple approfondi (Advanced Topics)

- Ce module étudiera de manière approfondie des sujets sélectionnés entourant le monde des « Data scientist ».

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours

Contenu du cours :

Table 1: KBP001 - Contenu du cours - Part A Python Development

Chapter	Description
Introduction	<ul style="list-style-type: none"> • Background • Why Python • Static vs. Managed Vs. Dynamic Languages • Usages • Development Tools
Python basic data types	<ul style="list-style-type: none"> • Expressions, Statements, Variables • Working with Numbers • Working with Strings • Working with Booleans • Working with Arrays • Input & Output
Python common operators	<ul style="list-style-type: none"> • If Statement • If-else statement • If-elif statement • Switch statement • While Loops • For Loops • Break & Continue • Range function
Collections	<ul style="list-style-type: none"> • Working with Lists • Working with Dictionaries • Working with Tuples • Working with Sets • Sorting Dictionaries • Copying Collections • Copying Collections – shallow and deep copy • Advanced usages with Map, Filter & Reduce

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours, Suite

Chapter	Description
Functions	<ul style="list-style-type: none">• Simple functions• Default Values• Named Args• Map• Filter• Reduce• Anonymous Functions• Built-In Functions• Scopes: Global , Local , Nested• Nested Functions• Lambda Expressions• Callback functions
Exceptions Handling	<ul style="list-style-type: none">• Bug, Error & Exception• Exception Syntax• Exception Levels• The try-catch-finally keywords• Custom Exceptions• Raising an Exception
Object Oriented Programming	<ul style="list-style-type: none">• Classes• Objects• Magic Functions• Effective documentation using docstrings• O.O. Concepts• Fields• Abstract Data Types• Inheritance• Special Methods• Virtual Functions

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours, Suite

Chapter	Description
Modules & Packages	<ul style="list-style-type: none"> • What are Modules ? • Importing Modules • Testing Modules • Built in Modules • User Defined Modules • Search Hierarchy • Multiple source files • Byte code • Packages & PyPi

Table 2: KBP001 - Contenu du cours - Part B Data Analysis with Python

Chapter	Description
Working with Files & Data	<ul style="list-style-type: none"> • Accessing files • Reading files • Writing to files • Text files • XML & JSON files • Binary Files • Yaml Files & ruaml_yaml package • Working with DB • Working with the Web
The NumPy Library	<ul style="list-style-type: none"> • Working with Array • Broadcasting • N-Dimensional Arrays • Reading Text Files
The Matplotlib Library	<ul style="list-style-type: none"> • Matplotlib objects • Plotting • The pyplot state machine

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours, Suite

Chapter	Description
The scipy Library	<ul style="list-style-type: none">• Useful usages• Non-Linear equations optimization• Non-Linear equations solvers• Root vs. Fsolve techniques
The Pandas Library (If time permits)	<ul style="list-style-type: none">• Series & Index• DataFrame• Group By

Table 3: KBP001 - Contenu du cours - Part C Machine Learning

Chapter	Description
Introduction to Machine Learning (ML)	<ul style="list-style-type: none">• The ML role on Artificial Intelligence world• Machine Learning vs. Deep Learning• ML Algorithms• Practical use cases• Methodologies & Tools• Scikit-Learn• Pandas• TensorFlow & PyTorch
ML Model lifecycle	<ul style="list-style-type: none">• Exploratory analysis• Data cleaning• Feature engineering• Algorithm selection• Model training• What's next?

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours, Suite

Chapter	Description
Data Preparation	<ul style="list-style-type: none"> • Data normalization • Data cleaning • Data filtering • Data scaling • Feature selection • Feature extraction
Supervised Learning Algorithms	<ul style="list-style-type: none"> • K-nearest neighbors • Linear models / Linear Regression • Logistic Regression • Decision trees • Random forest • Gradient boosting • Support vector machine (SVM) • Gradient descent • Validation & cross-validation • Over fitting & regularization • Error decomposition
Unsupervised Learning Algorithms	<ul style="list-style-type: none"> • Data dimensionality • Dimensionality reduction • Principal Component Analysis (PCA) • Kernel PCA • Centroid based clustering • K-means • K-medians • K-medoids • Density based clustering • DBSCAN • OPTICS • Distributing based clustering (Gaussian Mixture Model)

Ce sujet continue à la page suivante



KBP001 – Data Science Essentials, Suite

Contenu du cours, Suite

Table 4: KBP001 - Contenu du cours – Part D - Advanced Topics

Chapter	Description
Advanced Topics	<ul style="list-style-type: none">• Intro to Deep learning• Dimensionality reduction• Neural networks• Principal Component Analysis (PCA)• Kernel PCA• Intro to Recommendation Systems• Intro to Anomaly Detection
The End	<ul style="list-style-type: none">• Q&A• Course's Evaluation

