Company	SVM-TP-GPT		KNN-GPT	K-MEANS-GPT	Intelligence artificielle (IA) : L'intelligence artificielle	Composants d'un système expert :
Section Company Comp	2 import numpy as np	2	from sklearn.datasets import load_iris	2 import pandas as pd # pour la manipulation des données	de programmes et de machines capables de réaliser des	domaine d'application.
Company Comp	4 from sklearn.model_selection import		train_test_split	# pour charger l'ensemble de données Iris	humaine. Cela inclut des capacités telles que la	de connaissances pour tirer des conclusions ou prendre
Part	5 from sklearn.svm import SVC	5	<pre>from sklearn.neighbors import KNeighborsClassifier</pre>	# pour l'algorithme K-Means	visuelle, la prise de décision et la traduction entre	Interface utilisateur : Permet aux utilisateurs d'interagir
Section Company Comp	7	7			Espace d'états : Un espace d'états en intelligence	des réponses.
			autre jeu de données si nécessaire)	<pre>8 iris = load_iris()</pre>	représentation de tous les états possibles (configurations	à chaque cas traité par le système.
Company	12 X = data.drop('Class', axis=1)	10	X = iris.data	10	problème donné. Chaque action possible transforme un	les conclusions du système à l'utilisateur.
Company	uf 'Class'	12		12 kmeans = KMeans(n_clusters=3)	explorent cet espace pour trouver une solution au	 Recherche en profondeur (Depth-First Search, DFS) :
Company	14		et de test X train, X test, y train, y test = train test spli	13 kmeans.fit(X) # apprentissage du modèle 14		avant de revenir en arrière.
An international content of the co	et de test		t(X, y,	16 y_kmeans = kmeans.predict(X)		Explore tous les nœuds d'un niveau avant de passer au niveau suivant.
A continue of the continue o	train_test_split(X, y, test_size=0.2, random_state	15		18 # Affichage des clusters assignés à chaque échantillon 19 print("Clusters assignés à chaque échantillon :\n",	d'autres méthodes heuristiques ou optimisées selon le	coût le plus faible pour atteindre l'objectif, combinant
Company Comp	17 18 # Initialiser le classificateur SVM	18	X_train = sc.fit_transform(X_train)	20	Apprentissage supervisé : L'apprentissage supervisé est	coût uniforme.
Company	<pre>19 classifier = SVC(kernel='linear', random_state=42) 20</pre>	20	- '- '	22 print("\nCoordonnées des centres de clusters :\n",	est entraîné sur un jeu de données contenant des entrées	 Apprentissage supervisé: Les modèles prédisent des
Services and support and company to the company of		22	classifier = KNeighborsClassifier(n_neighbors=5)	23	modèle de prédire la sortie pour de nouvelles entrées.	d'algorithmes incluent la régression linéaire et les
Secretary of the secret		23		26 print("\nMatrice de confusion :\n", conf_matrix)	les prédictions continues et la classification pour les	 Apprentissage non supervisé: Les modèles identifient
Service of the control of the contro	26	25	classifier.fit(X_train, y_train)	· · · · · · · · · · · · · · · · · · ·	Apprentissage non supervisé : L'apprentissage non	Exemples d'algorithmes incluent K-means et l'analyse en
March	28 conf_matrix = confusion_matrix(y_test, y_pred)	27		2 import numpy as np 3 from sklearn.cluster import KMeans	où le modèle est entraîné sur un jeu de données sans	Apprentissage par renforcement : Les modèles
March Company Compan		29 30	# Calculer la précision du modèle	5	groupements ou les associations communes. Les	
March Marc	sklearn		print("Précision du modèle K-NN : {:.2f}%"		et l'analyse en composantes principales (ACP) pour la	 Régression: Utilisée pour prédire des valeurs continues.
Second Continue	3		KNN-YT		Régression, classification et regroupement :	sa localisation.
The content of the	y_pred والتوقعات test_	1	# Import des bibliothèques nécessaires	<pre>11 x.columns = ['Sepal_Length', 'Sepal_Width',</pre>	d'apprentissage supervisé visant à prédire des valeurs	discrètes. Exemple : déterminer si un email est un spam
The content of the	6		# pour les calculs numériques	12	en fonction de ses caractéristiques comme la	Business Intelligence (BI) : Technologie utilisée pour la
See			# pour la manipulation des données	r le nombre optimal de clusters	 Classification: La classification est une méthode 	destinées à soutenir les décisions d'affaires. Elle
West March Company			# pour les visualisations	15 for i in range(1, 11):	catégories discrètes (étiquettes). Par exemple,	
See The production of the control of			train_test_split	17 kmeans.fit(x)	 Regroupement (Clustering): Le clustering est une méthode d'apprentissage non supervisé utilisée pour 	Data Lifecycle Management (DLM): Gestion du flux de
1	ني مذا الكود، matplotlib.pyplot لم يتم استخدام # 13 لذلك لا داعي لاستيراده		aînement et de test	19	grouper un ensemble d'objets de manière que les objets dans le même groupe (ou cluster) soient plus	la suppression, avec des outils qui facilitent la gestion,
A first continue of the cont	14 # import matplotlib.pyplot as plt 15		KNeighborsClassifier # pour le modèle KNN	21 model = KMeans(n_clusters=3)	similaires (dans certains sens) les uns aux autres que ceux d'autres groupes. Par exemple, regrouper les	Cycle de vie des données :
March Section Company Compan	لتقسيم البيانات لتدريب واختبار النموذج rn		# pour charger les données de digits	24 # Matrice de confusion	habitudes d'achat.	 Stockage: Sécurisation des données collectées.
A continue of the continue o		0	# nour la matrice de confusion	26 matrice = confusion_matrix(iris.target, model.	l'apprentissage automatique sont des ensembles de	transformation des données.
AMAGEMENT CASES PARTICIPATION OF THE STATE OF THE CONTROL OF THE			# Chargement des données	labels_)	ou un label. Ces étiquettes servent de réponses correctes	interpréter les données.
Service of the content of the conten	21		# chargement des données		supervisé. Par exemple, dans un jeu de données pour la	
A second content of the property of the content o	لحساب مصفوفة الالتباس rn	13		import pandas as pd	le nom de l'objet qu'elle représente.	 analyses. Suppression: Élimination des données obsolètes.
Service of the control of the contro	24			from sklearn.datasets import load_iris	ensembles de données qui ne contiennent pas de labels	utilisé pour gérer et intégrer les données dans les
Proceedings	<pre>26 data = pd.read_csv('bill_authentication.csv') 27</pre>		X_train, X_test, y_train, y_test =	# pour l'algorithme K-Means	sont utilisées pour l'apprentissage non supervisé, où le	données de sources multiples, les transformant pour
Company Comp	train_test_split ر باستخدام			A pour le matrice de confusion	référence préalable.	 Différence entre ETL et ELT: L'ETL transforme les
The content of the	train_test_split(X, y, test_size=0.2)	18		# Chargement de l'ensemble de données Iris	technique de régression qui modélise la relation entre une	que l'ELT les transforme après le chargement.
Particular principal products products and products are producted by the control of the contro	SVM اِنشاء نعوذ # # 31		# initialisation du modèle avec k=5		données observées. Chaque observation est prédite	 Outils open source: Exemples incluent Talend et Pentaho
See Control (1) Co	33				Régression linéaire multiple : La régression linéaire	PowerCenter.
Particulation of continues Particulation	35 Classifier.fit(X_train, y_train)	22	# Calcul du taux de précision		qui utilise deux ou plus de variables indépendantes pour	Outils personnalisés : Développés pour répondre aux
## Septiment of the composition	يانات الاختبار		# evaluation du modele		ajuste une surface plane aux données, essayant de	 Types de sources de données : Bases de données, fichiers
and the control of th	39				réelles.	Lecture : Utilisation de composants comme
## State August Control Contro	التوقعات	26		print("Clusters assignés à chaque échantillon :\n",	méthode de classification utilisée pour prédire l'issue	ou des bases de données comme MySQL.
# All Productions and Contributions and Contribu	42	28	train_accuracy = np.empty(len(neighbors))		variables indépendantes. Elle modélise la probabilité que	lignes, conversion des types de données.
# First interest and processing and	sklearn لطباعة تقرير التصنيف		e d'entraînement	print((neodradinees des centres de ciusters . (n ,	spécifique.	de flux de données.
# print(content to content of the print of the of the prin	طباعة تقرير التمنيف باستخدام القيم الحقيقية و # 46		# initialisation de la précision sur l'ensembl 4	# Calcul de la macrice de Confusion	une donnée en analysant les K échantillons les plus	
# In section of the control of the c	47 print(classification_report(y_test, y_pred))		6		classification est souvent basée sur un vote majoritaire	
A contained and processes of the proce	1 # Importer les bibliothèques nécessaires				Arbres de décision : Un arbre de décision est un modèle	de données
Less de la martine processor in grove 1 Service (1997) de la martine (19	3 from sklearn.model_selection import			fonctionnement du cerveau humain, utilisés pour	chaque nœud interne représente un test sur un attribut,	données anormales.
# Chargement de domes (now pour course course) present au service de domes (now pour course course) per l'expertisse de domes et an experiment de l'experiment	4 from sklearn.preprocessing import StandardScaler	34	courante	les données. Ils sont composés de neurones disposés en	feuille représente une étiquette de classe (décision prise	redondantes. • Croisée des données: Comparaison de données
# Consparent des doubles (outs power utiliser un mire you de montes à indexession et a internation of the control of the contr	6 from sklearn.metrics import accuracy_score 7		# apprentissage du modèle	des calculs et transmet des sorties.	Forêt aléatoire : La forêt aléatoire est une méthode	 Harmonisation des valeurs et des dates: Standardisation
# ret_x_currexy() vm.rcurex(x_stx) y = ifs(x_streyt) y = ifs(x_s	autre jeu de données si nécessaire)		y_train)	lorsqu'un modèle d'apprentissage automatique est trop	décision lors de l'entraînement. Pour les prédictions, la	Codage et simplification: Transformation des données
3	9 iris = datasets.load_iris() 10 X = iris.data	36	test_accuracy[i] = knn.score(X_test,	nombre d'observations. Le modèle apprend les détails et le bruit des données d'entraînement au détriment de sa	moyenne des prédictions des arbres individuels (pour la régression) ou par le mode des classes (pour la	 Extraction des valeurs: Récupération des données
# a first test Security Secu	12			Sous-apprentissage : Le sous-apprentissage se produit	classification). Machine à vecteurs de support (SVM) : La SVM est une	 Modification de l'organisation des données:
trianglest, particle, particlest, particle, particlest, particlest	et de test		# affichage de la précision sur l'ensemble de	simple pour capturer la structure sous-jacente des	versatile utilisée pour la classification, la régression et la	meilleure gestion.
## Standardiser les caractéristiques en retirent 1 an noyeme et en nettrant 2 infection du mode s'article de confusion of se entres de s'article de confusion of se entres de confusion de se entres de	train_test_split(X, y, test_size=0.2, random_state			d'entraînement ni sur de nouvelles données, indiquant	maximise la marge entre les classes de données.	 tFileInputDelimited: Permet de lire un fichier texte délimité.
mayone et en nettant à 1'échaile la variance unit de la subtracé de confusion airs' sur profession partive ("est y pred) de confusion attrive ("est y pred) de confusion attrive ("est y pred) de confusion de coto que utilisate de mêthodes comme le direction de coto que utilisate de mêthodes comme le disposit que une comme le de policia participate. NN-DAILA NN-	15 16 # Standardiser les caractéristiques en retirant la	41		 Régularisation : La régularisation est une technique 	(clustering) consiste à partitionner un ensemble de	tMysqlinput: Permet de lire des données d'une base de
# matrice de confusion of print(confus) a strict, transform(C, train) print(confus) print(confus) print(confus) a strict set of confusion of print(confus) a strict set of confusion of print(confus) a strict set of confusion of confusion of print(confus) a strict set of confusion of confusio	moyenne et en mettant à l'échelle la variance unit aire	43		pénalisant la complexité du modèle. Elle aide à éviter le	groupe soient plus similaires entre elles qu'aux données	tOracleInput: Permet de lire des données d'une base de
# afficiage de la matrice de confusion # afficiage de la matrice de confusion # afficiage de la matrice de confusion # afficiateur 970 # afficiateur 970 # lombre d'occurrences de chaque classe dans y # test # monthe de conduction : une couche de convolution est une composante essembliel des réseaux de neurons # classifier - \$VIC(sernel-linear', random_statewa'2) # lombre d'occurrences de chaque classe dans y # test # monthe de conduction : une couche de convolution est une composante essembliel des réseaux de neurons # classifier - \$VIC(sernel-linear', random_statewa'2) # print("np. bin.count(y_test)) # MN-DARIJA # Sephdire las édiquettes pour l'ensemble de test # print ("np. bin.count (y_test)") # print("perfection du modèle # accuracy - accuracy - scuracy_score(y_test, y_pred) # accuracy - accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de confusion du modèle de laters # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de laters # accuracy - scuracy_score(y_test, y_pred) # # sinch matrice de laters # accuracy - scuracy_score(y_test, y_pred) # # accuracy - scuracy_s	18 X_train = sc.fit_transform(X_train)		<pre># matrice de confusion print(cm)</pre>	fonction de coût ou en utilisant des méthodes comme le dropout pour les réseaux de neurones.	techniques d'apprentissage non supervisé et aide à découvrir la structure inhérente des données sans	Filtrer et convertir les données
22 de Sassifiar - SVC(cerneles*Inager*, random_state=42) 28	20 21 # Initialiser le classificateur SVM	46	# affichage de la matrice de confusion	Couche de convolution : Une couche de convolution est une composante essentielle des réseaux de neurones	étiquettes préalables. Regroupement hiérarchique : Le regroupement	données.
threather is models easy (lassifier, mytrain) predict les étiquettes pour l'ensemble de test production models est et de l'assifiéer, predict(X,test) prediction de models est et de l'assifiéer, predict(X,test) curacy * 100)) from sklean-datasets import load_iris else hichéado b database d library from sklean-nedplose import WeighborsClassifier accuracy * 100)) from sklean-nedplose import WeighborsClassifier accuracy * 100) from sklean-nedplose import WeighborsClassifier accuracy * 100)) from sklean-nedplose import WeighborsClassifier accuracy * 100) from sklean-nedplose import WeighborsClassifie	<pre>22 classifier = SVC(kernel='linear', random_state=42) 23</pre>		_test	traitement d'image. Elle effectue une opération	construire une hiérarchie de clusters. Elle peut être	
# Prédire les étiquettes pour l'ensemble de test 1 From sitleann, database d'althrary		48		avec un ensemble de poids, appelé un filtre ou noyau,	progressivement les clusters les plus proches) ou divisive	d'une colonne.
2 from sklearm.nedel_selection import train_test_plit which neems data li jendna from sklearm.nedel_selection import train_test_plit which neems data li jendna from sklearm.nedel_selection import darkindedo b nodel knn whichedo b nodel knn from sklearm.nedel_selection import scurvey * 10e)) 2 from sklearm.nedel_selection import train_test_plit which neems data li jendna from sklearm.nedel_selection import darkindedo b nodel knn whichedo b nodel knn from sklearm.nedel_selection import darkindedo b nodel knn whichedo b accuracy_score black inhiedo b accuracy_score black information stell black informations les plus courament utilises and the unispert darkitation les plus courament at unities ean punitises as soft end minimisant a variance intra-cluster et annialismant a variance intra-cluster et annialismant avariance intra-cluster et annialismant avar			from sklearn.datasets import load_iris	Fonction d'activation : La fonction d'activation est une	permettant ainsi de visualiser des structures de données	données. • Agréger et joindre les données
31 accuracy = accuracy_score(y_text, y_pred) 32 print("Precision du modèle SWI : (:.2f)%", format(4 print("precision du modèle SWI : (:.2f)%", format(5 print("precision du modèle SWI : (:.2f)%", format(6 print("precision du modèle SWI : (:.2f)%", format(7 print("precision du modèle SWI : (:.2f)%", format(8 print("precision du modèle SWI : (:.2f)%", format(8 print("precision du modèle SWI : (:.2f)%", format(9 print("precision du modèle SWI : (:.2f)%", format(12 print("precision du modèle SWI : (:.2f)%", format(13 print("precision du modèle SWI : (:.2f)%", format(14 print("precision du modèle SWI : (:.2f)%", format(15 print("precision du modèle SWI : (:.2f)%", format(15 print("precision du modèle SWI : (:.2f)%", format(16 print("precision du modèle SWI : (:.2f)%", format(17 print("precision du modèle SWI : (:.2f)%", format(18 print("precision du modè	29	2	from sklearn.model_selection import train_test_split #bach nqesmo data li 3endna	dans un réseau de neurones. ReLU est l'une des fonctions •	Regroupement par partition : Le regroupement par	 tAggregateRow: Permet d'agréger les données d'un flux de données.
**Couche de pooling: Une couche de pooling ou de sau-matrix confusion baccuracy w ndrio matrix confusion firs = load_riss() **Inàja khas njibo data firs = load_riss() **Radio hishedmo baccuracy w ndrio matrix confusion of ## suban hishedmo baccuracy w ndrio matrix confusion of ## suban hishedmo baccuracy w ndrio matrix confusion of ## suban hishedmo baccuracy and provided and service services of ## suban hishedmo baccuracy accuracy = accuracy = couche gooling: Une couche de pooling ou de sous-échantillonnage réduit la dimensionnalité state of earactéristiques tout en conservant les informations les plus importantes. Elle est souvent utilisée de laractéristiques tout en conservant les informations les plus importantes. Elle est souvent utilisée de la ralaction et al dimensionnalité : La réduction de dimension divis en analyse de données dans une baccuracy de sindifférents formats une visual réduire le nombre de paramètres et le sur-apprentissage : La réduction de dimension div	31 accuracy = accuracy_score(y_test, y_pred) 32 print("Précision du modèle SVM : {:.2f}%".format(3	<pre>from sklearn.neighbors import KNeighborsClassifier #bach nkhedmo b model knn</pre>	les valeurs d'entrée positives sans changement et annule	des données en plusieurs groupes (partitions) en minimisant la variance intra-cluster et en maximisant la	triées d'un flux de données.
de chaque carte de caractéristiques tout en conservant les informations les plus importantes. Elle est souvent utilisée dans les réseaux de neurones convolutifs pour réduire le nombre de paramètres et le sur-apprentissage. 8			#bach nkhedmo b accuracy w ndiro matrix confusion	Couche de pooling: Une couche de pooling ou de sous-échantillonnage réduit la dimensionnalité spatiale	variance inter-cluster. Chaque groupe contient des données plus similaires entre elles qu'aux données des	tMap: Permet de mapper les données d'un flux de
uninser dars de la principales. Calcular per médiure le nombre de paramètres et le sur-apprentissage automatique et et analyse de données qui visé a réduire le nombre de paramètres et le sur-apprentissage. 3		6	iris = load_iris()	de chaque carte de caractéristiques tout en conservant les informations les plus importantes. Elle est souvent	autres groupes. Réduction de dimensionnalité : La réduction de	 Écrire les données dans différents formats
**Strain, X_test, Y_train, Y_test = Considérée, par obtention d'un ensemble de variables principales. Ceta permet de simplifier les modèles et de réduire le coût computationnel tout en préservant le plus possible l'information utile. 14 An sayob londel s 3 tina 1 k = 3 15 k = 3 16 Mn = *Kleighborsclassifier(n_neighbors=k) 17 Mn = Kleighborsclassifier(n_neighbors=k) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit(X_train, Y_train) 15 Mn.fit(X_train, Y_train) 16 Mn.fit(X_train, Y_train) 17 Mn.fit(X_train, Y_train) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit(X_train, Y_train) 15 Mn.fit(X_train, Y_train) 16 Mn.fit(X_train, Y_train) 17 Mn.fit(X_train, Y_train) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit(X_train, Y_train) 15 Mn.fit(X_train, Y_train) 16 Mn.fit(X_train, Y_train) 17 Mn.fit(X_train) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train) 10 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit(X_train, Y_train) 15 Mn.fit(X_train, Y_train) 16 Mn.fit(X_train, Y_train) 17 Mn.fit(X_train, Y_train) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit(X_train, Y_train) 15 Mn.fit(X_train, Y_train) 16 Mn.fit(X_train, Y_train) 17 Mn.fit(X_train, Y_train) 18 Mn.fit(X_train, Y_train) 19 Mn.fit(X_train, Y_train) 10 Mn.fit(X_train, Y_train) 11 Mn.fit(X_train, Y_train) 12 Mn.fit(X_train, Y_train) 13 Mn.fit(X_train, Y_train) 14 Mn.fit		8	X, y = iris.data, iris.target	utilisée dans les réseaux de neurones convolutifs pour	dimensionnalité est un processus en apprentissage automatique et en analyse de données qui vise à réduire	base de données SQL Server.
### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### An anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Anaybo Inodel s 3tina 1 k = 3 ### Liapprentissage profond Liapprentissage profond sun assure profonds (composed set on an anaybo Inodel set 1 tilloutputtSon: Permet d'écrire des données dans un dichier Excel. ### Liapprentissage automatique qui utilise des réseaux de neurones profonds (composés de nombreuses couches) pour modéliser des relations complexes entre les entrées et découvir des motifs dans les données de grande dimension. ### An anaybo Inodel s 3tina 1 k = 3 ### L'apprentissage profond : Lapprentissage profond et un sous-domnées dans un dichier Excel. ### Liapprentissage profond sun unities. ### L'apprentissage profond sun unities. ### Couche Inode Reservative (a controller a voice et decouvir des nombreuses couches) pour modéliser des relations complexes entre les entrées et découvir des nombreuses couches pour modéliser des relations complexes entre les entrées et découvir des nombreuses couches pour modéliser des relations complexes entre les entrées et découvir des nombreuses couches pour modéliser des relations complexes entre les entrées et découvir des nombreuses couches pour modéliser des relations complexes et découvir des nombreuses couches pour modéliser des relations complexes et découvir des nombreuses couches pour modéliser des relations complexes et decouvir des nombreuses couches pour modéliser des relations complexes et découvir des nombreuses couches pour modéliser des relations complexes et decouvir des nombreuses couches pour modéliser des relations complexes et découvir des nombreuses couches pour modéliser des relations complexes et dé		10	X_train, X_test, y_train, y_test =		considérées, par obtention d'un ensemble de variables	
12 k = 3 13 km = KleighborsClassifier(n_neighbors=k) 14 #5 entrink 1 module 15 km.fit(X_train, y_train) 16 #6 nkhedno b predict 17 y_pred = km.n_predict(X_test) 18 #7 nkhedno b accuracy 19 accuracy = accuracy_score(y_test, y_pred) 20 print(f^Accuracy): {accuracy} or accuracy brint(fracuracy) 21 print(f^Accuracy): {accuracy} 22 print(f^Accuracy): {accuracy} 23 print(f^Accuracy): {accuracy} 24 print(f^Accuracy): {accuracy} 25 print(f^Accuracy): {accuracy} 26 print(f^Accuracy): {accuracy} 27 print(f^Accuracy): {accuracy} 28 print(f^Accuracy): {accuracy} 29 print(f^Accuracy): {accuracy} 20 print(f^Accuracy): {accuracy} 21 print(f^Accuracy): {accuracy} 22 print(f^Accuracy): {accuracy} 23 print(f^Accuracy): {accuracy} 24 print(f^Accuracy): {accuracy} 25 print(f^Accuracy): {accuracy} 26 print(f^Accuracy): {accuracy} 27 print(f^Accuracy): {accuracy} 28 print(f^Accuracy): {accuracy} 29 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 31 print(f^Accuracy): {accuracy} 32 print(f^Accuracy): {accuracy} 33 print(f^Accuracy): {accuracy} 34 print(f^Accuracy): {accuracy} 35 print(f^Accuracy): {accuracy} 36 print(f^Accuracy): {accuracy} 37 print(f^Accuracy): {accuracy} 38 print(f^Accuracy): {accuracy} 39 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 31 print(f^Accuracy): {accuracy} 32 print(f^Accuracy): {accuracy} 33 print(f^Accuracy): {accuracy} 34 print(f^Accuracy): {accuracy} 35 print(f^Accuracy): {accuracy} 36 print(f^Accuracy): {accuracy} 37 print(f^Accuracy): {accuracy} 37 print(f^Accuracy): {accuracy} 38 print(f^Accuracy): {accuracy} 39 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 30 print(f^Accuracy): {accuracy} 31 print(f^Accuracy): {accuracy} 32 print(f^Accuracy): {accuracy} 33 print(f^Accuracy): {accuracy}			=42)		réduire le coût computationnel tout en préservant le plus	un fichier texte délimité. • tFileOutputExcel: Permet d'écrire des données dans un
14 #5 entrinis I module 15 km.nfit/\tanhan, \tanhan, \tan		12 13	k = 3 knn = KNeighborsClassifier(n_neighbors=k)		L'apprentissage profond : L'apprentissage profond est un	tFileOutputJSON: Permet d'écrire des données dans un
16 #6 nkbedmo b predict 17 y_pred = knn.predict(X_test) 18 #7 nkbedmo b accuracy 19 accuracy = accuracy_score(y_test, y_pred) 20 print(f*Accuracy: {accuracy}s(x)) 21 print(f*Accuracy: {accuracy}s(x)) 22 print(f*Accuracy: {accuracy}s(x)) 23 print(f*Accuracy: {accuracy}s(x)) 24 print(f*Accuracy: {accuracy}s(x))		14 15	#5 entriniw l module knn.fit(X_train, y_train)		des réseaux de neurones profonds (composés de	tichier JSON.
18 #7 nichedmo b accuracy 19 accuracy - accuracy - accuracy - Couche entièrement connectée : Une couche 20 print(f"accuracy: {accuracy}-sore) 21 print(f"accuracy: {accuracy}-sore) 22 print(f"accuracy: {accuracy}-sore) 23 print(f"accuracy: {accuracy}-sore) 24 print(f"accuracy: {accuracy}-sore) 25 print(f"accuracy: {accuracy}-sore) 26 print(f"accuracy: {accuracy}-sore) 27 print(f"accuracy: {accuracy}-sore) 28 print(f"accuracy: {accuracy}-sore) 39 print(f"accuracy: {accuracy}-sore) 30 print(f"accuracy: {accuracy}-sore) 30 print(f"accuracy: {accuracy}-sore) 31 print(f"accuracy: {accuracy}-sore) 32 print(f"accuracy: {accuracy}-sore) 33 print(f"accuracy: {accuracy}-sore) 34 print(f"accuracy: {accuracy}-sore) 35 print(f"accuracy: {accuracy}-sore) 36 print(f"accuracy: {accuracy}-sore) 37 print(f"accuracy: {accuracy}-sore) 38 print(f"accuracy: {accuracy}-sore) 39 print(f"accuracy: {accuracy}-sore) 30 p		17	y_pred = knn.predict(X_test)		complexes entre les entrées et les sorties et découvrir des motifs dans les données de grande dimension.	
20 print (+"Accuracy: {accuracy}") où chaque neurone est connecté à tous les neurones de la		19	accuracy = accuracy_score(y_test, y_pred)	•	Couche entièrement connectée : Une couche entièrement connectée, ou couche dense, est une couche	
		20	print(T Accuracy: {accuracy}")		où chaque neurone est connecté à tous les neurones de la couche précédente. Cette couche combine les	
caractéristiques apprises pour classer						

Partie 1 : questions de cours . Donne le nom de trois composants Talend appartenant à la phase Extraction :

d'une colonn

base de données Oracle

un fichier texte délimité

d'autres systèmes ou applications

les lignes uniques

-tConvertType Transformer les do

Partie 3 : Étude de cas

Partie 2 : Liaison par flèche Lier chaque composant Talend avec son rôle. -tFileInputDelimited Lecture d'un fichier CSV -tUniqueRow Permet de détecter les lignes redondantes et de garder

nnées en modifiant leur type de

Sélectionner des lignes de données vérifiant des critères de sélection -tMysqlOutput Charger un contenu dans une BD MySQL

tFileInputDelimited : Permet de lire un fichier texte

tFileInputExcel: Permet de lire un fichier Excel.
tMysqlinput: Permet de lire des données d'une base de
données MySQL.

tDenormalizeSortedRow: Permet de dénormaliser les

tOracleOutput : Permet d'écrire des données dans une

Pourquoi on utilise le composant tLogRow? Le composant tLogRow est utilisé pour afficher le contenu des données traitées dans la console ou les logs, ce qui permet de vérifier et de déboguer les flux de données pendant l'exécution.

la phase du processus ETL où les données brutes sont nettoyées et converties aux formats appropriés est :
 [x] Transformation

Correction Control 1 ETL
 le rôle principal de l'ETL dans la BI:
 [x] Faciliter la collecte, la transformation et le chargement

Donne le nom de trois composants Talend appartenant à • Le Data Lifecycle Management (DLM): la phase Transformation : • [x] Un processus de gestion des données du début à la fin tritlerColumns: Permet de filtrer les colonnes d'un flux

Les composants clés d'un processus ETL :
 [x] Extraction, Transformation, Chargement

tConvertType : Permet de convertir le type de données L'objectif des outils de restitution : [x] Générer des rapports et des tableaux de bord

Donne le nom de trois composants Talend appartenant à la phase Chargement : tMSSqlOutput: Permet d'écrire des données dans une base de données SQL Server.

L'extraction/chargement incrémental(e) : [x] ETL ne charge que les données modifiées/ajoutées depuis la dernière exécution. tFileOutputDelimited : Permet d'écrire des données dans

Où est positionné l'ETL dans la chaîne décisionnelle :

Que signifie l'acronyme OLAP ? [x] Online Analytical Processing La différence entre un Datamart et un Data Warehouse : pendant texecution.

Pourquoi on utilise le composant tFileOutputDelimited?

Le composant tFileOutputDelimited est utilisé pour écrire des données dans un fichier texte délimité, permettant ainsi d'exporter et de sauvegarder les données traitées dans un format standardisé et facilement lisible par d'autres pur believes que policier.

La Business Intelligence (BI), ou informatique décisionnelle, est la technologie utilisée pour la collecte, l'analyse et la production d'informations pour soutenir les décisions d'affaires.
 Un système d'information décisionnel est structuré pour transformer les données en informations utiles pour

prendre des décisions stratégiques. Il utilise des processus ETL pour intégrer et préparer les données,

données.

I'ETL est composé ? L'ETL est composé de trois phases :

Etraction, Transformation, Chargement. Mul06, page 31.

Donne quatre types de sources de données ?

Bases de données

Fichiers plats

Services web ERP/CRM

Convertir le type de Prix_Unitaire de string à Double (1

Faire une agrégation des Prix totaux des produits par Ville

tAggregateRow

Faire un tri ascendant par Prix totaux des produits

tConvertType
Faire une jointure des deux fichiers

Faire un affichage des résultats sur la console sous forme de tableau

tLogRow

My Un Data Warehouse contient toutes les données de l'entreprise, tandis qu'un Datamart contient des données spécifiques à un domaine

processus ETL pour intégrer et préparer les données, stockées ensuite dans des entrepôts de données et analysées via des outils de reporting. Un système d'information opérationnel gère les opérations quotidiennes d'une organisation, incluant des ERP pour la gestion des resouveres, CRM pour la gestion des clients, et autres bases de données pour stocker les données transactionnelles. L'ETL (Extract, Transform, Load) est un processus en trois d'anne parmetant d'extraire des données transers de dranse parmetant d'extraire des données transers de

étapes permettant d'extraire des données brutes de sources multiples, de les transformer pour les préparer à l'utilisation, et de les charger dans un entrepôt de

Quelle est la différence entre données structurées et non

structurées?
Les données structurées sont organisées selon un modèle prédéfini comme des bases de données relationnelles, tandis que les données non structurées ne suivent pas de modèle spécifique, comme les emails, les vidéos, les Fichiers texte. MLIGb. _age_11
Donne trois composants permettant l'extraction des dronnées dans Talend?

tFileInputDelimited : Permet de lire un fichier texte

délimité.

##HielnputExcel : Permet de lire un fichier Excel.

##Mysqlinput : Permet de lire des données d'une base de données MySQL.

Donne trois exemples d'outils ETL populaires ?

Talend Open Studio
Informatica PowerCenter
Microsoft SSIS

Dessine l'architecture détaillée d'une chaîne

Dessine l'architecture détaillée d'une chaîne décisionnelle d'un projet Bl et expliquer les rôles des couches de l'architecture ? : Collecte des données : Etraction des données des systèmes sources. Transformation des données : Nettoyage, standardisation et consolidation des données : Nettoyage, standardisation et consolidation des données : Tengistrement des données transformées données : Italiation des données transformées dans un Data Warehouse.

Restitution des données : Utilisation des données via des

outils de reporting et d'analyse pour soutenir la prise de