

SVM-TP-GPT		KNN-GPT		K-MEANS-GPT		Intelligence artificielle (IA) : L'intelligence artificielle désigne la discipline scientifique consacrée à la création de programmes et de machines capables de réaliser des tâches qu, traditionnellement, nécessitent l'intelligence humaine. Cela inclut des capacités telles que la compréhension du langage naturel, la reconnaissance visuelle, la prise de décision et la traduction entre langues.		Composants d'un système expert :	
1	Importez les bibliothèques nécessaires	1	Import des bibliothèques nécessaires	1	import numpy as np # pour les calculs numériques	2		Base de connaissances : Contient les faits et les règles du domaine d'application.	
2	import numpy as np	2	from sklearn.datasets import load_iris	2	import pandas as pd # pour la manipulation des données	3		Moteur d'inférence : Traite les règles et les faits de la base de connaissances pour tirer des conclusions ou prendre des décisions.	
3	from sklearn.model_selection import train_test_split	3	from sklearn.model_selection import train_test_split	3	from sklearn.datasets import load_iris	4		Interface utilisateur : Permet aux utilisateurs d'interagir avec le système expert, posant des questions et recevant des réponses.	
4	from sklearn.preprocessing import StandardScaler	4	from sklearn.metrics import confusion_matrix	4	from sklearn.metrics import confusion_matrix	5		Base de données de faits : Stocke les données spécifiques à chaque cas traité par le système.	
5	from sklearn.metrics import accuracy_score	5	from sklearn.metrics import accuracy_score	5	from sklearn.metrics import accuracy_score	6		Mécanisme d'explication : Explique les raisonnements et les conclusions du système à l'utilisateur.	
6	from sklearn.metrics import confusion_matrix	6	from sklearn.metrics import confusion_matrix	6	from sklearn.metrics import confusion_matrix	7		Trois algorithmes de recherche :	
7		7		7		8		Recherche en profondeur (Depth-First Search, DFS) :	
8	Charger les données à partir du fichier CSV	8	Chargement des données (vous pouvez utiliser un autre jeu de données si nécessaire)	8	Chargement de l'ensemble de données Iris	9		Explore autour que possible le long de chaque branche avant de revenir en arrière.	
9	data = pd.read_csv('bill_authentication.csv')	9	iris = load_iris()	9	iris = load_iris()	10		Recherche en largeur (Breadth-First Search, BFS) :	
10		10	X = iris.data	10	X = iris.data	11		Explore tous les nœuds d'un niveau avant de passer au niveau suivant.	
11	Diviser les caractéristiques et les étiquettes	11	Y = iris.target	11	Y = iris.target	12		Recherche A* : Utilise des heuristiques pour estimer le coût le plus faible pour atteindre l'objectif, combinant ainsi les approches de meilleure première recherche et de coût uniforme.	
12	X = data.drop('Class', axis=1)	12		12		13		Types d'apprentissage automatique :	
13	Les caractéristiques sont toutes les colonnes sauf 'Class'	13	Diviser les données en ensembles d'entraînement et de test	13	Diviser les données en ensembles d'entraînement et de test	14		Apprentissage supervisé : Les modèles prédisent des sorties basées sur des entrées étiquetées. Exemples d'algorithmes incluent la régression linéaire et les machines à vecteurs de support (SVM).	
14	y = data['Class'] # 'Class' est notre étiquette	14	X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)	14	X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)	15		Apprentissage non supervisé : Les modèles identifient des structures cachées dans des données non étiquetées. Exemples d'algorithmes incluent K-means et l'analyse en composantes principales (PCA).	
15	Diviser les données en ensembles d'entraînement et de test	15		15		16		Apprentissage par renforcement : Les modèles apprennent à prendre des décisions en observant les conséquences de leurs actions, sans données étiquetées pré-fournies. Un exemple d'algorithme est Q-learning.	
16	X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)	16	Standardisation des caractéristiques	16	Standardisation des caractéristiques	17		Différence entre régression et classification :	
17		17	sc = StandardScaler()	17	sc = StandardScaler()	18		Régression : Utilisée pour prédire des valeurs continues. Exemple : prédire le prix d'une maison basé sur sa taille et sa localisation.	
18	Initialiser le classificateur SVM	18	X_train = sc.fit_transform(X_train)	18	X_train = sc.fit_transform(X_train)	19		Classification : Utilisée pour prédire des catégories discrètes. Exemple : déterminer si un email est un spam ou non.	
19	classifier = SVC(kernel='linear', random_state=42)	19	X_test = sc.transform(X_test)	19	X_test = sc.transform(X_test)	20		Business Intelligence (BI) : Technologie utilisée pour la collecte, l'analyse et la production d'informations destinées à soutenir les décisions d'affaires. Elle transforme les données en informations exploitables et aide les entreprises à comprendre et à agir sur leurs données pour prendre les meilleures décisions.	
20		20	Initialiser le classificateur K-NN	20	Initialiser le classificateur K-NN	21		Data Lifecycle Management (DLM) : Gestion du flux de données tout au long de leur cycle de vie, de la collecte à la suppression, avec des outils qui facilitent la gestion, réduisent les coûts et automatisent les étapes du cycle.	
21	classifier.fit(X_train, y_train)	21	classifier = KNeighborsClassifier(n_neighbors=5)	21	classifier = KNeighborsClassifier(n_neighbors=5)	22		Cycle de vie des données :	
22		22	Spécifiez le nombre de voisins (k)	22	Spécifiez le nombre de voisins (k)	23		Collecte : Acquisition et saisie de nouvelles données.	
23		23		23		24		Stockage : Sécurisation des données collectées.	
24	Prédire les étiquettes pour l'ensemble de test	24	y_pred = classifier.predict(X_test)	24	y_pred = classifier.predict(X_test)	25		Traitement : Compression, cryptage, nettoyage, et transformation des données.	
25	y_pred = classifier.predict(X_test)	25		25		26		Analyse : Utilisation de techniques d'analyse pour interpréter les données.	
26	Calculer la matrice de confusion	26		26		27		Sauvegarde : Stockage sécurisé des données sur différents supports.	
27	conf_matrix = confusion_matrix(y_test, y_pred)	27		27		28		Réutilisation : Utilisation des données pour de nouvelles analyses.	
28	print("Matrice de confusion :\n", conf_matrix)	28		28		29		Suppression : Élimination des données obsolètes.	
SVM-TP		KNN-YT		RANDOM FOREST		Non étiquetées : Les données non étiquetées sont des ensembles de données qui ne contiennent pas de labels ou d'étiquettes indiquant la sortie désirée. Ces données sont utilisées pour l'apprentissage non supervisé, où le système doit identifier des motifs ou des structures sans référence préalable.		Types d'outils ETL :	
1	من مكتبة classification_report اختيار الدالة sklearn	1	Import des bibliothèques nécessaires	1	import numpy as np # pour les calculs numériques	2		Outils open source : Exemples incluent Talend et Pentaho.	
2	from sklearn.metrics import classification_report	2	import numpy as np	2	import pandas as pd	3		Logiciels d'entreprise : Exemple inclut Informatica PowerCenter.	
3	from sklearn.metrics import classification_report	3	import pandas as np	3	import pandas as np	4		Outils basés sur le cloud : Exemple inclut Amazon Glue.	
4	from sklearn.metrics import classification_report	4	import matplotlib.pyplot as plt	4	import matplotlib.pyplot as plt	5		Outils personnalisés : Développés pour répondre aux besoins spécifiques d'une entreprise.	
5	from sklearn.metrics import classification_report	5	from sklearn.model_selection import train_test_split	5	from sklearn.model_selection import train_test_split	6		Types de sources de données : Bases de données, fichiers plats, services web.	
6	from sklearn.metrics import classification_report	6	from sklearn.metrics import classification_report	6	from sklearn.metrics import classification_report	7		Manipulation des données :	
7	from sklearn.metrics import classification_report	7	from sklearn.metrics import classification_report	7	from sklearn.metrics import classification_report	8		Lecture : Utilisation de composants comme TFileInputDelimited pour lire des fichiers texte délimités ou des bases de données comme MySQL.	
8	from sklearn.metrics import classification_report	8	from sklearn.metrics import classification_report	8	from sklearn.metrics import classification_report	9		Filtrage et conversion : Filtrage des colonnes ou des lignes, conversion des types de données.	
9	from sklearn.metrics import classification_report	9	from sklearn.metrics import classification_report	9	from sklearn.metrics import classification_report	10		Agrégation et jonction : Agrégation des données, jonction de flux de données.	
10	from sklearn.metrics import classification_report	10	from sklearn.metrics import classification_report	10	from sklearn.metrics import classification_report	11		Écriture : Écriture des données dans différents formats et bases de données.	
11	from sklearn.metrics import classification_report	11	from sklearn.metrics import classification_report	11	from sklearn.metrics import classification_report	12		Types de données	
12	from sklearn.metrics import classification_report	12	from sklearn.metrics import classification_report	12	from sklearn.metrics import classification_report	13		Valeurs manquantes : Données non présentes dans le jeu de données	
13	from sklearn.metrics import classification_report	13	from sklearn.metrics import classification_report	13	from sklearn.metrics import classification_report	14		Détection des valeurs aberrantes : Identification des données anormales.	
14	from sklearn.metrics import classification_report	14	from sklearn.metrics import classification_report	14	from sklearn.metrics import classification_report	15		Élimination des doublons : Suppression des données redondantes.	
15	from sklearn.metrics import classification_report	15	from sklearn.metrics import classification_report	15	from sklearn.metrics import classification_report	16		Croisement des données : Comparaison de données provenant de différentes sources.	
16	from sklearn.metrics import classification_report	16	from sklearn.metrics import classification_report	16	from sklearn.metrics import classification_report	17		Harmonisation des valeurs et des dates : Standardisation des formats de données.	
17	from sklearn.metrics import classification_report	17	from sklearn.metrics import classification_report	17	from sklearn.metrics import classification_report	18		Codage et simplification : Transformation des données pour simplifier leur usage.	
18	from sklearn.metrics import classification_report	18	from sklearn.metrics import classification_report	18	from sklearn.metrics import classification_report	19		Extraction des valeurs : Récupération des données pertinentes.	
19	from sklearn.metrics import classification_report	19	from sklearn.metrics import classification_report	19	from sklearn.metrics import classification_report	20		Modification de l'organisation des données : Réorganisation des structures de données pour une meilleure gestion.	
20	from sklearn.metrics import classification_report	20	from sklearn.metrics import classification_report	20	from sklearn.metrics import classification_report	21		Lire les données à partir de différentes sources	
21	from sklearn.metrics import classification_report	21	from sklearn.metrics import classification_report	21	from sklearn.metrics import classification_report	22		TFileInputDelimited : Permet de lire un fichier texte délimité.	
22	from sklearn.metrics import classification_report	22	from sklearn.metrics import classification_report	22	from sklearn.metrics import classification_report	23		TFileInputExcel : Permet de lire un fichier Excel.	
23	from sklearn.metrics import classification_report	23	from sklearn.metrics import classification_report	23	from sklearn.metrics import classification_report	24		TMySQLinput : Permet de lire des données d'une base de données MySQL.	
24	from sklearn.metrics import classification_report	24	from sklearn.metrics import classification_report	24	from sklearn.metrics import classification_report	25		TOracleinput : Permet de lire des données d'une base de données Oracle.	
25	from sklearn.metrics import classification_report	25	from sklearn.metrics import classification_report	25	from sklearn.metrics import classification_report	26		Filter et convertir les données	
26	from sklearn.metrics import classification_report	26	from sklearn.metrics import classification_report	26	from sklearn.metrics import classification_report	27		TFilterColumns : Permet de filtrer les colonnes d'un flux de données.	
27	from sklearn.metrics import classification_report	27	from sklearn.metrics import classification_report	27	from sklearn.metrics import classification_report	28		TConvertType : Permet de convertir le type de données d'une colonne.	
28	from sklearn.metrics import classification_report	28	from sklearn.metrics import classification_report	28	from sklearn.metrics import classification_report	29		TDNormalizeSortedRow : Permet de dénormaliser les données.	
29	from sklearn.metrics import classification_report	29	from sklearn.metrics import classification_report	29	from sklearn.metrics import classification_report	30		TAgrégateRow : Permet d'agréger les données d'un flux de données.	
30	from sklearn.metrics import classification_report	30	from sklearn.metrics import classification_report	30	from sklearn.metrics import classification_report	31		TJoin : Permet de joindre deux flux de données.	
31	from sklearn.metrics import classification_report	31	from sklearn.metrics import classification_report	31	from sklearn.metrics import classification_report	32		TMap : Permet de mapper les données d'un flux de données.	
32	from sklearn.metrics import classification_report	32	from sklearn.metrics import classification_report	32	from sklearn.metrics import classification_report	33		TWrite : Permet d'écrire des données dans différents formats	
33	from sklearn.metrics import classification_report	33	from sklearn.metrics import classification_report	33	from sklearn.metrics import classification_report	34		TMSQOutput : Permet d'écrire des données dans une base de données SQL Server.	
34	from sklearn.metrics import classification_report	34	from sklearn.metrics import classification_report	34	from sklearn.metrics import classification_report	35		TOracleOutput : Permet d'écrire des données dans une base de données Oracle.	
35	from sklearn.metrics import classification_report	35	from sklearn.metrics import classification_report	35	from sklearn.metrics import classification_report	36		TFileOutputDelimited : Permet d'écrire des données dans un fichier texte délimité.	
36	from sklearn.metrics import classification_report	36	from sklearn.metrics import classification_report	36	from sklearn.metrics import classification_report	37		TFileOutputExcel : Permet d'écrire des données dans un fichier Excel.	
37	from sklearn.metrics import classification_report	37	from sklearn.metrics import classification_report	37	from sklearn.metrics import classification_report	38		TFileOutputJSON : Permet d'écrire des données dans un fichier JSON.	
38	from sklearn.metrics import classification_report	38	from sklearn.metrics import classification_report	38	from sklearn.metrics import classification_report	39			
39	from sklearn.metrics import classification_report	39	from sklearn.metrics import classification_report	39	from sklearn.metrics import classification_report	40			
40	from sklearn.metrics import classification_report	40	from sklearn.metrics import classification_report	40	from sklearn.metrics import classification_report	41			
41	from sklearn.metrics import classification_report	41	from sklearn.metrics import classification_report	41	from sklearn.metrics import classification_report	42			
42	from sklearn.metrics import classification_report	42	from sklearn.metrics import classification_report	42	from sklearn.metrics import classification_report	43			
43	from sklearn.metrics import classification_report	43	from sklearn.metrics import classification_report	43	from sklearn.metrics import classification_report	44			
44	from sklearn.metrics import classification_report	44	from sklearn.metrics import classification_report	44	from sklearn.metrics import classification_report	45			
45	from sklearn.metrics import classification_report	45	from sklearn.metrics import classification_report	45	from sklearn.metrics import classification_report	46			
46	from sklearn.metrics import classification_report	46	from sklearn.metrics import classification_report	46	from sklearn.metrics import classification_report	47			
47	from sklearn.metrics import classification_report	47	from sklearn.metrics import classification_report	47	from sklearn.metrics import classification_report	48			
48	from sklearn.metrics import classification_report	48	from sklearn.metrics import classification_report	48	from sklearn.metrics import classification_report	49			
49	from sklearn.metrics import classification_report	49	from sklearn.metrics import classification_report	49	from sklearn.metrics import classification_report	50			
50	from sklearn.metrics import classification_report	50	from sklearn.metrics import classification_report	50	from sklearn.metrics import classification_report	51			
51	from sklearn.metrics import classification_report	51	from sklearn.metrics import classification_report	51	from sklearn.metrics import classification_report	52			
52	from sklearn.metrics import classification_report	52	from sklearn.metrics import classification_report	52	from sklearn.metrics import classification_report	53			
53	from sklearn.metrics import classification_report	53	from sklearn.metrics import classification_report	53	from sklearn.metrics import classification_report	54			
54	from sklearn.metrics import classification_report	54	from sklearn.metrics import classification_report	54	from sklearn.metrics import classification_report	55			
55	from sklearn.metrics import classification_report	55	from sklearn.metrics import classification_report	55	from sklearn.metrics import classification_report	56			
56	from sklearn.metrics import classification_report	56	from sklearn.metrics import classification_report	56	from sklearn.metrics import classification_report	57			
57	from sklearn.metrics import classification_report	57	from sklearn.metrics import classification_report	57	from sklearn.metrics import classification_report	58			
58	from sklearn.metrics import classification_report	58	from sklearn.metrics import classification_report	58	from sklearn.metrics import classification_report	59			
59	from sklearn.metrics import classification_report	59	from sklearn.metrics import classification_report	59	from sklearn.metrics import classification_report	60			
60	from sklearn.metrics import classification_report	60	from sklearn.metrics import classification_report	60	from sklearn.metrics import classification_report	61			
61	from sklearn.metrics import classification_report	61	from sklearn.metrics import classification_report	61	from sklearn.metrics import classification_report	62			
62	from sklearn.metrics import classification_report	62	from sklearn.metrics import classification_report	62	from sklearn.metrics import classification_report	63			
63	from sklearn.metrics import classification_report	63	from sklearn.metrics import classification_report	63	from sklearn.metrics import classification_report	64			
64	from sklearn.metrics import classification_report	64	from sklearn.metrics import classification_report	64	from sklearn.metrics import classification_report	65			
65	from sklearn.metrics import classification_report	65	from sklearn.metrics import classification_report	65	from sklearn.metrics import classification_report	66			
66	from sklearn.metrics import classification_report	66	from sklearn.metrics import classification_report	66	from sklearn.metrics import classification_report	67			
67	from sklearn.metrics import classification_report	67	from sklearn.metrics import classification_report	67	from sklearn.metrics import classification_report	68			
68	from sklearn.metrics import classification_report	68	from sklearn.metrics import classification_report	68	from sklearn.metrics import classification_report	69			
69	from sklearn.metrics import classification_report	69	from sklearn.metrics import classification_report	69	from sklearn.metrics import classification_report	70			
70	from sklearn.metrics import classification_report	70	from sklearn.metrics import classification_report	70	from sklearn.metrics import classification_report	71			
71	from sklearn.metrics import classification_report	71	from sklearn.metrics import classification_report	71	from sklearn.metrics import classification_report	72			
72	from sklearn.metrics import classification_report	72	from sklearn.metrics import classification_report	72	from sklearn.metrics import classification_report	73			
73	from sklearn.metrics import classification_report	73	from sklearn.metrics import classification_report	73	from sklearn.metrics import classification_report	74			
74	from sklearn.metrics import classification_report	74	from sklearn.metrics import classification_report	74	from sklearn.metrics import classification_report	75			
75	from sklearn.metrics import classification_report	75	from sklearn.metrics import classification_report	75	from sklearn.metrics import classification_report	76			
76	from sklearn.metrics import classification_report	76	from sklearn.metrics import classification_report	76	from sklearn.metrics import classification_report	77			
77	from sklearn.metrics import classification_report	77	from sklearn.metrics import classification_report	77	from sklearn.metrics import classification_report	78			
78	from sklearn.metrics import classification_report	78	from sklearn.metrics import classification_report	78	from sklearn.metrics import classification_report	79			
79	from sklearn.metrics import classification_report	79	from sklearn.metrics import classification_report	79	from sklearn.metrics import classification_report	80			
80	from sklearn.metrics import classification_report	80	from sklearn.metrics import classification_report	80	from sklearn.metrics import classification_report	81			
81	from sklearn.metrics import classification_report	81	from sklearn.metrics import classification_report	81	from sklearn.metrics import classification_report	82			
82	from sklearn.metrics import classification_report	82	from sklearn.metrics import classification_report	82	from sklearn.metrics import classification_report	83			
83	from sklearn.metrics import classification_report	83	from sklearn.metrics import classification_report	83	from sklearn.metrics import classification_report	84			
84	from sklearn.metrics import classification_report	84	from sklearn.metrics import classification_report	84	from sklearn.metrics import classification_report	85			
85	from sklearn.metrics import classification_report	85	from sklearn.metrics import classification_report	85	from sklearn.metrics import classification_report	86			
86	from sklearn.metrics import classification_report	86	from sklearn.metrics import classification_report	86	from sklearn.metrics import classification_report	87			
87	from sklearn.metrics import classification_report	87	from sklearn.metrics import classification_report	87	from sklearn.metrics import classification_report	88			
88	from sklearn.metrics import classification_report	88	from sklearn.metrics import classification_report	88	from sklearn.metrics import classification_report	89			
89	from sklearn.metrics import classification_report	89	from sklearn.metrics import classification_report	89	from sklearn.metrics import classification_report	90			
90	from sklearn.metrics import classification_report	90	from sklearn.metrics import classification_report	90	from sklearn.metrics import classification_report	91			
91	from sklearn.metrics import classification_report	91	from sklearn.metrics import classification_report	91	from sklearn.metrics import classification_report	92			
92	from sklearn.metrics import classification_report	92	from sklearn.metrics import classification_report	92	from sklearn.metrics import classification_report	93			
93	from sklearn.metrics import classification_report	93	from sklearn.metrics import classification_report	93	from sklearn.metrics import classification_report	94			
94	from sklearn.metrics import classification_report	94	from sklearn.metrics import classification_report	94	from sklearn.metrics import classification_report	95			
95	from sklearn.metrics import classification_report	95	from sklearn.metrics import classification_report	95	from sklearn.metrics import classification_report	96			
96	from sklearn.metrics import classification_report	96	from sklearn.metrics import classification_report	96	from sklearn.metrics import classification_report	97			
97	from sklearn.metrics import classification_report	97	from sklearn.metrics import classification_report	97	from sklearn.metrics import classification_report	98			
98	from sklearn.metrics import classification_report	98	from sklearn.metrics import classification_report	98	from sklearn.metrics import classification_report	99			
99	from sklearn.metrics import classification_report	99	from sklearn.metrics import classification_report	99	from sklearn.metrics import classification_report	100			
100	from sklearn.metrics import classification_report	100	from sklearn.metrics import classification_report	100	from sklearn.metrics import classification_report	101			
101	from sklearn.metrics import classification_report	101	from sklearn.metrics import classification_report	101	from sklearn.metrics import classification_report	102			
102	from sklearn.metrics import classification_report	102	from sklearn.metrics import classification_report	102	from sklearn.metrics import classification_report	103			
103	from sklearn.metrics import classification_report	103	from sklearn.metrics import classification_report	103	from sklearn.metrics import classification_report	104			
104	from sklearn.metrics import classification_report	104	from sklearn.metrics import classification_report	104	from sklearn.metrics import classification_report	105			
105	from sklearn.metrics import classification_report	105	from sklearn.metrics import classification_report	105	from sklearn.metrics import classification_report	106			
106	from sklearn.metrics import classification_report	106	from sklearn.metrics import classification_report	106	from sklearn.metrics import classification_report	107			
107	from sklearn.metrics import classification_report	107	from sklearn.metrics import classification_report	107	from sklearn.metrics import classification_report	108			
108	from sklearn.metrics import classification_report	108	from sklearn.metrics import classification_report	108	from sklearn.metrics import classification_report	109			
109	from sklearn.metrics import classification_report	109	from sklearn.metrics import classification_report	109	from sklearn.metrics import classification_report	110			
110	from sklearn.metrics import classification_report	110	from sklearn.metrics import						

- **Partie 1 : questions de cours .**
- **Donne le nom de trois composants Talend appartenant à la phase Extraction :**
- **tFileInputDelimited :** Permet de lire un fichier texte délimité.
- **tFileInputExcel :** Permet de lire un fichier Excel.
- **tMySQLinput :** Permet de lire des données d'une base de données MySQL.
- **Correction Control 1 ETL**
- **le rôle principal de l'ETL dans la BI :**
- **[d] Faciliter la collecte, la transformation et le chargement des données**
- **la phase du processus ETL où les données brutes sont nettoyées et converties aux formats appropriés est :**
- **[x] Transformation**
- **Donne le nom de trois composants Talend appartenant à la phase Transformation :**
- **tFilterColumns :** Permet de filtrer les colonnes d'un flux de données.
- **tConvertType :** Permet de convertir le type de données d'une colonne.
- **tDenormalizeSortedRow :** Permet de dénormaliser les données.
- **Le Data Lifecycle Management (DLM) :**
- **[d] Un processus de gestion des données du début à la fin**
- **Les composants clés d'un processus ETL :**
- **[d] Extraction, Transformation, Chargement**
- **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.
- **tOracleOutput :** Permet d'écrire des données dans une base de données Oracle.
- **tFileOutputDelimited :** Permet d'écrire des données dans un fichier texte délimité.
- **Le fonctionnement de Talend repose sur :**
- **[d] L'installation de JDK et l'utilisation du langage Java**
- **L'extraction/chargement incrémental(e) :**
- **[d] ETL ne charge que les données modifiées/ajoutées depuis la dernière exécution.**
- **Où est positionné l'ETL dans la chaîne décisionnelle :**
- **[x] En amont**
- **Que signifie l'acronyme OLAP ?**
- **[d] Online Analytical Processing**
- **La différence entre un Datamart et un Data Warehouse :**
- **[d] Un Data Warehouse contient toutes les données de l'entreprise, tandis qu'un Datamart contient des données spécifiques à un domaine**
- **Questions de cours**
- **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, 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 ressources, 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 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 données.
- **L'ETL est composé ?** L'ETL est composé de trois phases : Extraction, Transformation, Chargement. [M106, page 31](#)
- **Donne quatre types de sources de données ?**
- Bases de données
- Fichiers plats
- Services web
- ERP/CRM
- **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. [M106, page 11](#)
- **Donne trois composants permettant l'extraction des données dans Talend ?**
- **tFileInputDelimited :** Permet de lire un fichier texte délimité.
- **tFileInputExcel :** Permet de lire un fichier Excel.
- **tMySQLinput :** Permet de lire des données d'une base de données MySQL.
- **Donne trois exemples d'outils ETL populaires ?**
- Talend Open Studio
- Informatica Power Center
- Microsoft SSIS
- **Dessine l'architecture détaillée d'une chaîne décisionnelle d'un projet BI et explique les rôles des couches de l'architecture ? :**
- **Collecte des données :** Extraction des données des systèmes sources.
- **Transformation des données :** Nettoyage, standardisation et consolidation des données.
- **Stockage des données :** Enregistrement 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 décision.