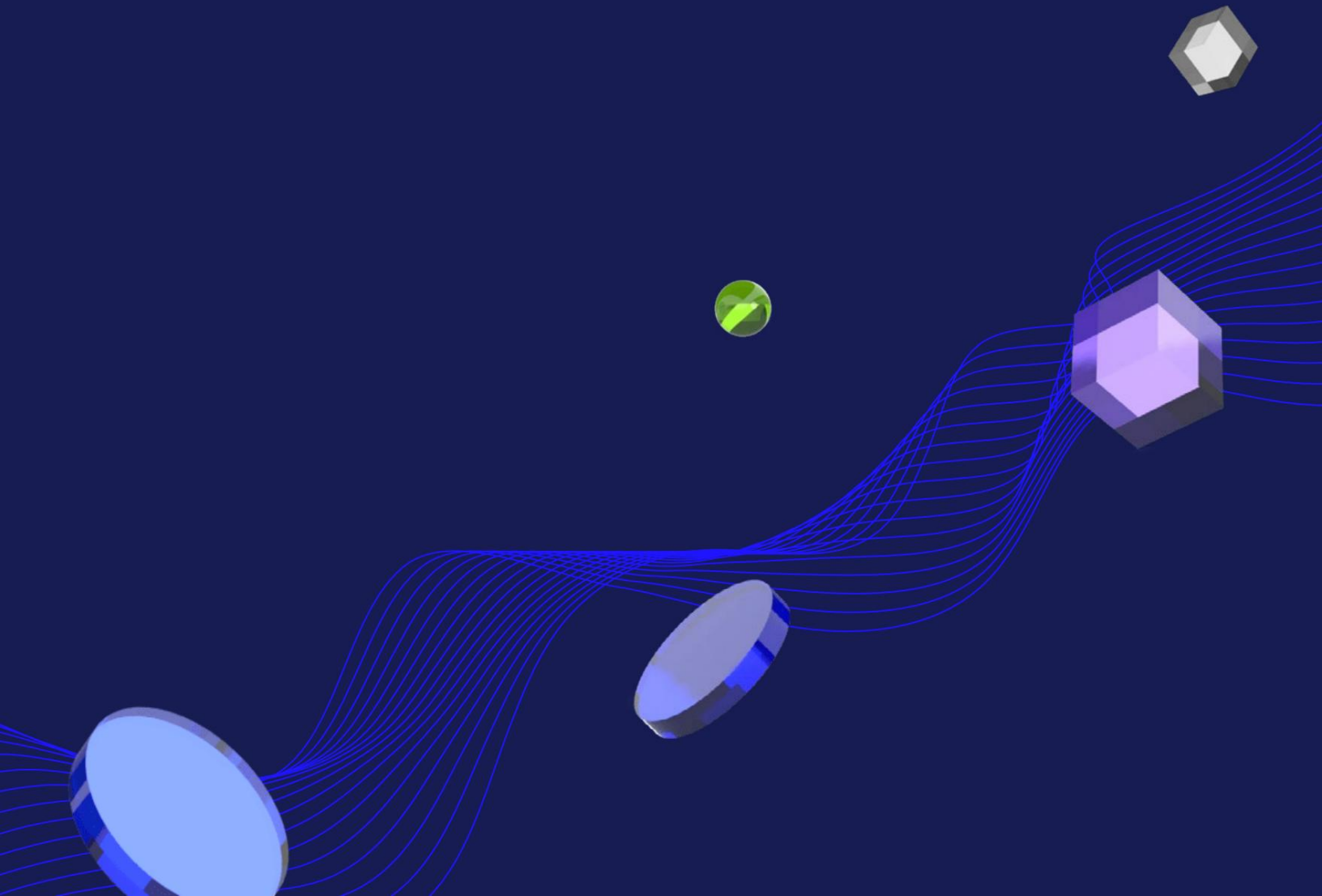




ÉCOLE D'INTELLIGENCE ARTIFICIELLE

# Langage naturel Traitement

Programme du programme Nanodegree



# Aperçu

Ce programme permettra d'améliorer les compétences existantes des apprenants en apprentissage automatique et en apprentissage profond grâce à l'ajout de techniques de traitement du langage naturel et de reconnaissance vocale. Ces compétences peuvent être utilisées dans diverses applications telles que l'étiquetage des parties du discours et la traduction automatique, entre autres. Les apprenants développeront les compétences dont ils ont besoin pour commencer à appliquer les techniques de traitement du langage naturel aux défis et aux applications du monde réel.

Réalisé en collaboration avec :



# Informations sur le programme



## Durée estimée

3 mois à raison de 10-15h/semaine\*



Niveau de compétence

Avancé



## Prérequis

Un apprenant bien préparé doit avoir une expérience significative de Python et une expérience de base en probabilités, statistiques et architectures d'apprentissage profond. Les apprenants doivent également avoir la capacité d'écrire une classe en Python et d'ajouter des commentaires à leur code pour que d'autres puissent le lire. Enfin, les apprenants doivent être familiarisés avec le terme « réseaux neuronaux » et les mathématiques différentielles qui conduisent à la rétropropagation.



## Matériel/logiciel requis

Les apprenants doivent avoir accès à un système d'exploitation 64 bits avec au moins 8 Go de RAM, ainsi qu'à des autorisations de compte administrateur suffisantes pour installer des programmes, notamment Anaconda avec Python 3.5 et les packages de support. Le réseau doit permettre une connexion sécurisée connexions à des hôtes distants (comme SSH).

\*La durée de ce programme est une estimation du nombre total d'heures que l'étudiant moyen peut prendre pour terminer tous les cours obligatoires, y compris les cours magistraux et les projets. Si vous consacrez environ 5 à 10 heures par semaine à suivre le programme, vous devriez le terminer dans le délai prévu. Le nombre d'heures réel peut varier.

# Introduction au traitement du langage naturel

Apprenez les principes fondamentaux du traitement de texte, notamment la recherche de radicaux et la lemmatisation. Explorez les méthodes d'apprentissage automatique dans l'analyse des sentiments. Créez un modèle d'étiquetage vocal.



## Projet de cours

### Balisateur des parties du discours

Utilisez plusieurs techniques, notamment les recherches dans les tables, les n-grammes et les modèles de Markov cachés, pour étiqueter les parties du discours dans les phrases et comparez leurs performances. Ce projet démontre des techniques de traitement de texte qui permettent de créer un modèle d'étiquetage des parties du discours. Travaillez avec une table de recherche simple et ajoutez progressivement plus de complexité pour améliorer le modèle à l'aide de modèles graphiques probabilistes. Utilisez un package Python pour créer et entraîner un étiqueteur avec un modèle de Markov caché et comparez les performances de tous ces modèles dans un ensemble de données de phrases.

#### Leçon 1

Introduction à la PNL

- Apprendre les principales techniques utilisées dans le traitement du langage naturel.
- Se familiariser avec la terminologie et les sujets abordés en classe.
- Construire sa première application avec IBM Watson.

#### Leçon 2

Traitement de texte

- Voir comment le texte est traité afin de l'utiliser dans les modèles.
- Apprenez des techniques telles que la tokenisation, la recherche de radicaux et la lemmatisation.
- Commencez avec l'étiquetage des parties du discours et la reconnaissance des entités nommées.

### Leçon 3

#### Balises des parties du discours avec Modèles de Markov cachés

- Apprenez comment les modèles de Markov cachés sont définis.
- Former les HMM avec les algorithmes de Viterbi et de Baum-Welch.
- Utiliser des HMM pour créer un modèle d'étiquetage des parties du discours.

### Cours 2

# Calculer avec le langage naturel

Apprenez des techniques avancées telles que l'intégration de mots, l'attention par apprentissage profond et bien plus encore. Créez un modèle de traduction automatique à l'aide d'architectures de réseaux neuronaux récurrents.



### Projet de cours

## Traduction automatique

Créez un réseau neuronal profond qui fonctionne dans le cadre d'un pipeline de traduction automatique de bout en bout. Le pipeline terminé acceptera le texte anglais en entrée et renverra la traduction française. Être capable d'explorer plusieurs architectures de réseaux neuronaux récurrents et de comparer leurs performances. Prétraiter les données en convertissant le texte en séquence d'entiers. Construisez plusieurs modèles d'apprentissage profond pour traduire le texte en français. Exécutez ces modèles sur un test anglais pour analyser leurs performances.

### Leçon 1

Extraction de caractéristiques

& Incorporations

- Apprenez à extraire des fonctionnalités d'un texte.
- Apprenez les algorithmes d'intégration les plus utilisés, tels que Word2Vec et Glove.

### Leçon 2

Modélisation

- Découvrez les principales utilisations des modèles d'apprentissage profond en PNL.
- Apprenez-en plus sur la traduction automatique, les modèles de sujets et l'analyse des sentiments.

### Leçon 3

Apprentissage profond Attention

- Découvrez l'attention, la méthode avancée d'apprentissage en profondeur qui permet d'utiliser des applications comme Google Translate.
- Apprenez-en plus sur l'attention additive et multiplicative dans des applications telles que la machine traduction, résumé de texte et légende d'images.
- Découvrez des modèles d'apprentissage profond de pointe comme le transformateur qui étendent l'utilisation de l'attention pour éliminer le besoin de RNN.

### Leçon 4

Systèmes d'information

- Apprendre les systèmes d'extraction et de recherche d'informations.
- Apprenez-en plus sur les réponses aux questions et leurs applications.

Cours 3

# Communiquer avec le langage naturel

Apprenez les techniques d'interface utilisateur vocale qui transforment la parole en texte et vice versa. Créez un modèle de reconnaissance vocale à l'aide de la reconnaissance vocale en profondeur réseaux neuronaux.



## Projet de cours

# Reconnaissance vocale

Créez un réseau neuronal profond qui fonctionne dans le cadre d'un pipeline de reconnaissance vocale automatique de bout en bout (ASR). Le modèle convertira l'audio brut en représentations de fonctionnalités, qui les transformeront ensuite en texte transcrit. Commencez par étudier un ensemble de données qui sera utilisé pour former et évaluer les modèles.

Convertissez n'importe quel fichier audio brut en représentations de fonctionnalités couramment utilisées pour la reconnaissance automatique de texte. Créez des réseaux neuronaux qui associent ces fonctionnalités au texte transcrit.

## Leçon 1

### Introduction aux interfaces utilisateur vocales

- Apprenez les bases de la façon dont les ordinateurs comprennent les mots parlés.
- Familiarisez-vous avec les applications VUI les plus courantes.
- Configurez un compte AWS et développez la compétence Alexa avec un modèle existant.

## Leçon 2

### Compétence Histoire d'Alexa

- Apprenez les bases d'Amazon AWS.
- Créez votre propre compétence Alexa entièrement fonctionnelle à l'aide de l'API d'Amazon.
- Déployer sa compétence pour que chacun puisse l'utiliser.

## Leçon 3

### Introduction à Reconnaissance vocale

- Apprenez le pipeline utilisé pour la reconnaissance vocale.
- Apprendre à traiter et à extraire des caractéristiques des signaux sonores.
- Apprenez à construire des modèles de langage probabilistes et d'apprentissage automatique afin de extraire des mots et de la grammaire à partir de signaux sonores.

# Rencontrez vos instructeurs.



**Luis Serrano**

Instructeur

Luis était auparavant ingénieur en apprentissage automatique chez Google. Il est titulaire d'un doctorat en mathématiques de l'Université du Michigan et d'une bourse postdoctorale de l'Université du Québec à Montréal. Montréal.



**Jay Alammar**

Directeur des investissements chez STV

Jay est diplômé en informatique, adore visualiser les concepts d'apprentissage automatique et est le directeur des investissements chez STV, un fonds de capital-risque de 500 millions de dollars axé sur la haute technologie. startups.



**Arpan Chakraborty**

Informaticien

Arpan est un informaticien titulaire d'un doctorat de l'Université d'État de Caroline du Nord. Il enseigne à Georgia Tech (dans le cadre du programme de maîtrise en informatique) et est co-auteur du livre Practical Graph Mining with R.



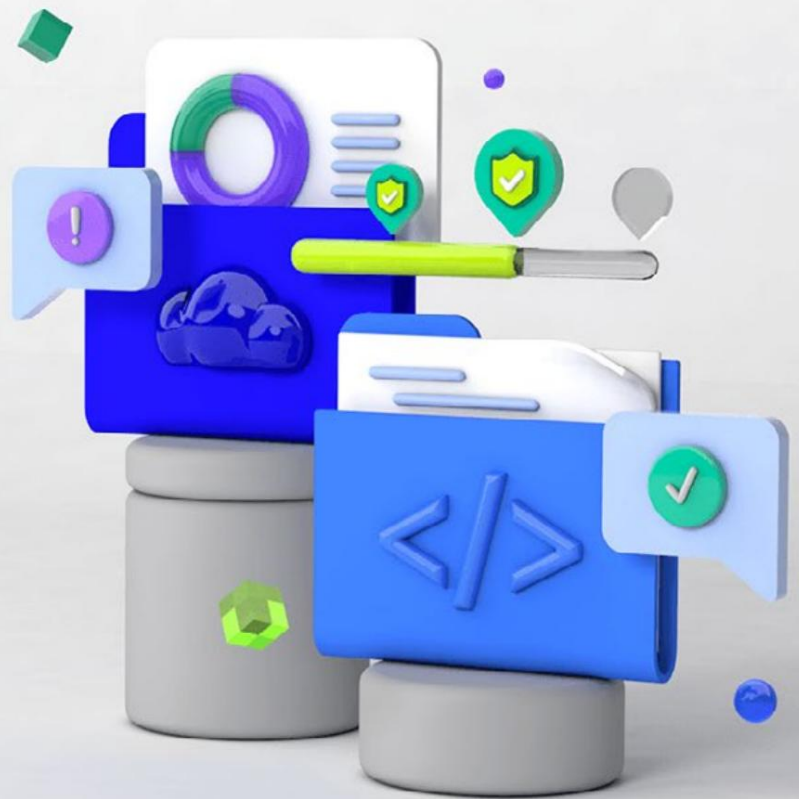
**Dana Sheahen**

Ingénieur électricien

Dana est ingénieure électricienne et titulaire d'un master en informatique de Georgia Tech. Son expérience professionnelle comprend le développement de logiciels pour les systèmes embarqués au sein du groupe automobile de Motorola, où elle a obtenu un brevet pour un système d'exploitation embarqué.



## L'expérience d'apprentissage d'Udacity



### Projets pratiques

Les projets expérimentaux ouverts sont conçus pour refléter les défis réels du milieu de travail. Il ne s'agit pas simplement de questions à choix multiples ou de guides étape par étape, mais plutôt d'une réflexion critique.



### Connaissance

Trouvez des réponses à vos questions avec Knowledge, notre wiki propriétaire. Recherchez les questions posées par d'autres étudiants, se connecter avec des mentors techniques, et découvrez comment résoudre les défis qui vous rencontrent.



### Espaces de travail

Visualisez votre code en action. Vérifiez le résultat et la qualité de votre code en l'exécutant sur des espaces de travail interactifs intégrés à la plateforme.



### Quiz

Les quiz auto-notés renforcent la compréhension.

Les apprenants peuvent revenir aux leçons à tout moment pendant le cours pour rafraîchir les concepts.



### Programmes d'études personnalisés

Créez un plan d'études personnalisé qui répond à vos besoins individuels. Utilisez ce plan pour suivre vos progrès vers votre objectif global.

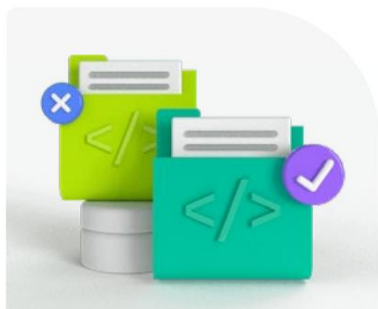


### Suivi des progrès

Profitez des rappels d'étapes clés pour rester dans les temps et terminer votre programme.

## Notre approche éprouvée pour la construction

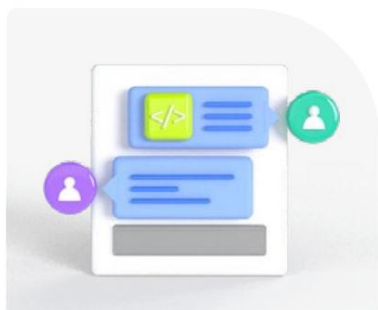
### Compétences numériques prêtes à l'emploi.



#### Examineurs de projets expérimentés

Vérifier la maîtrise des compétences.

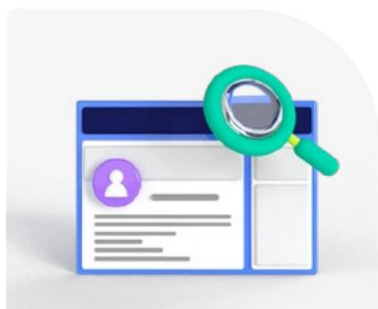
- Les commentaires et critiques personnalisés sur le projet incluent un examen du code ligne par ligne par des praticiens qualifiés avec un délai d'exécution moyen de 1,1 heure.
- Le cycle de révision du projet crée une boucle de rétroaction avec de multiples opportunités d'amélioration – jusqu'à ce que le concept soit maîtrisé.
- Les examinateurs de projets s'appuient sur les meilleures pratiques du secteur et fournissent des conseils de pro.



#### Soutien du mentor technique

L'assistance 24h/24 et 7j/7 débloque l'apprentissage.

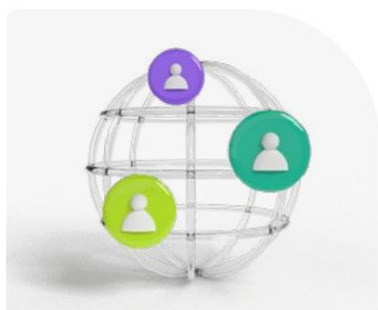
- L'apprentissage s'accélère à mesure que des mentors qualifiés identifient les domaines de réussite et de potentiel pour la croissance.
- L'accès illimité aux mentors signifie que l'aide arrive quand elle est le plus nécessaire.
- Un temps de réponse moyen aux questions de 2 heures ou moins garantit que le développement des compétences reste sur la bonne voie.



#### Services de carrière personnels

Favoriser la préparation à l'emploi.

- Accès à une revue de portefeuille Github qui peut vous donner un avantage en mettant en valeur vos points forts et démontrer votre valeur aux employeurs.\*
- Obtenez de l'aide pour optimiser votre LinkedIn et établir votre marque personnelle afin que votre profil est mieux classé dans les recherches effectuées par les recruteurs et les responsables du recrutement.



#### Réseau de mentorat

Hautement contrôlé pour son efficacité.

- Les mentors doivent suivre un processus de recrutement en 5 étapes pour rejoindre le réseau sélectif d'Udacity.
- Après avoir réussi une évaluation objective et situationnelle, les mentors doivent démontrer communication et adéquation comportementale pour un rôle de mentorat.
- Les mentors travaillent dans plus de 30 secteurs différents et suivent souvent eux-mêmes un programme Nanodegree.

\*S'applique uniquement à certains programmes Nanodegree.



Pour en savoir plus, rendez-vous sur

[www.udacity.com/online-learning-for-individuals](https://www.udacity.com/online-learning-for-individuals) →