

# **Unidad N°1 Roadmap recomendado para continuar tu transformación hacia el rol de un profesional experto en Data Science**

Sitio: [Centro de E-Learning - UTN.BA](#)

Curso: Curso de Data Science

Libro: Unidad N°1 Roadmap recomendado para continuar  
tu transformación hacia el rol de un profesional  
experto en Data Science

Imprimido  
por:

Virginia Marich

Día:

Wednesday, 27 de December de 2023,  
09:53

## Descripción

## Tabla de contenidos

1. Introducción
2. Roadmap
3. Introducción a GitHub para la creación de repositorios
4. Temáticas avanzadas a profundizar y bibliografía recomendada
5. Conclusiones

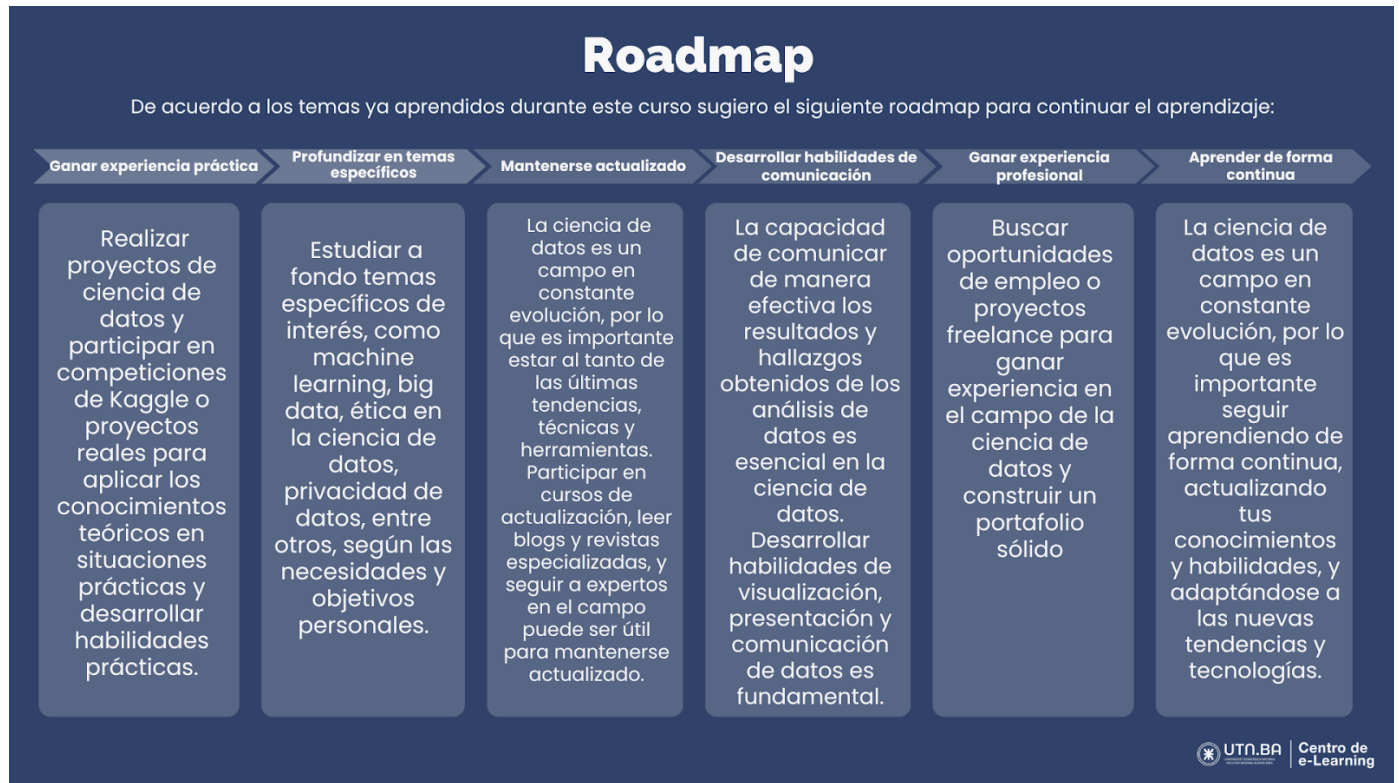
# 1. Introducción

¡Bienvenido/a a la continuación de este curso de ciencia de datos! Ahora que ya hemos cubierto varios temas importantes, es momento de enfocarnos en cómo seguir aprendiendo y desarrollándote en este emocionante campo. En esta unidad, te proporcionaremos un roadmap para que puedas planificar tu progreso en la ciencia de datos de manera efectiva.

En esta unidad también te introduciremos a GitHub, una plataforma ampliamente utilizada para la creación de repositorios y la colaboración en proyectos de desarrollo de software. Aprender a utilizar GitHub te permitirá tener una herramienta adicional para gestionar tus proyectos de ciencia de datos y trabajar de manera colaborativa en un entorno profesional.

## 2. Roadmap

De acuerdo a los temas ya aprendidos durante este curso sugiero el siguiente roadmap para continuar el aprendizaje:



### 3. Introducción a GitHub para la creación de repositorios

Si eres un estudiante reciente en el campo de la ciencia de datos, construir un portafolio en GitHub puede ser una excelente manera de demostrar tus habilidades y experiencia en el campo. Aquí hay algunos pasos que puedes seguir para crear un portafolio efectivo en GitHub:

1. Crea una cuenta en GitHub: Ve a <https://github.com> y haz clic en el botón verde "Sign up" para registrarte en GitHub de forma gratuita. Completa los campos requeridos, como tu nombre de usuario, dirección de correo electrónico y contraseña, y sigue las instrucciones para crear tu cuenta.
2. Verifica tu dirección de correo electrónico: Después de registrarte, recibirás un correo electrónico de verificación en la dirección de correo electrónico que proporcionaste. Haz clic en el enlace de verificación en el correo electrónico para activar tu cuenta de GitHub.
3. Inicia sesión en GitHub: Una vez que hayas verificado tu dirección de correo electrónico, inicia sesión en GitHub utilizando tu nombre de usuario y contraseña.
4. Crea un nuevo repositorio: Después de iniciar sesión, estarás en tu perfil de GitHub. Haz clic en el botón verde "New" en la esquina superior derecha para crear un nuevo repositorio.
5. Nombre tu repositorio: En la página de creación de repositorio, ingresa un nombre descriptivo para tu repositorio. Puedes usar el nombre de tu proyecto de ciencia de datos o algo que describa claramente el contenido de tu repositorio.
6. Establece la visibilidad del repositorio: Puedes elegir si quieres que tu repositorio sea público o privado. Si quieres que tu portafolio sea visible para cualquier persona, selecciona "Public", si prefieres que solo ciertas personas puedan acceder a él, selecciona "Private". Ten en cuenta que los repositorios privados requieren una suscripción paga en GitHub.
7. Agrega un archivo README: Un archivo README es una descripción del contenido y propósito de tu repositorio. Puedes agregar un archivo README marcando la casilla "Add a README file" en la página de creación de repositorio. Esto te permitirá proporcionar información adicional sobre tu proyecto y ayudará a los visitantes a entender lo que has hecho.
8. Crea el repositorio: Haz clic en el botón verde "Create repository" para crear tu repositorio en GitHub.
9. Descarga e instala Git: Git es una herramienta de control de versiones que necesitarás para interactuar con tu repositorio en GitHub. Puedes descargar Git desde <https://git-scm.com/> y seguir las instrucciones de instalación para tu sistema operativo.
10. Clona el repositorio en tu computadora: Una vez que hayas instalado Git, necesitarás clonar tu repositorio en tu computadora local para poder trabajar en él. Para hacer esto, abre una ventana de terminal o línea de comandos en tu computadora y navega a la ubicación donde deseas clonar tu repositorio. Luego, ejecuta el siguiente comando:

```
git clone URL_DEL_REPOSITORIO
```

Reemplaza "URL\_DEL\_REPOSITORIO" con la URL de tu repositorio en GitHub. Esto creará una copia local de tu repositorio en tu computadora.

11. Agrega tus archivos y carpetas: Ahora puedes agregar los archivos y carpetas de tu proyecto de ciencia de datos al repositorio clonado en tu computadora local. Puedes hacer esto mediante la línea de comandos o utilizando una herramienta de control de versión

Disponible en <https://youtu.be/uM6evmCAW8U>

Disponible en <https://youtu.be/3xKpVmseEzI>

## 4. Temáticas avanzadas a profundizar y bibliografía recomendada

1. Aprendizaje Profundo (Deep Reinforcement Learning): Esta es una rama avanzada del aprendizaje automático que combina redes neuronales profundas con algoritmos de aprendizaje por refuerzo. Aprender sobre cómo entrenar modelos de aprendizaje profundo para tomar decisiones secuenciales y aprender de forma autónoma a través de la retroalimentación del entorno, te permitirá abordar problemas complejos como el control de robots, la toma de decisiones en tiempo real y la optimización de políticas.

- "Spinning Up in Deep RL" de OpenAI:

<https://spinningup.openai.com/en/latest/>

Este es un conjunto de recursos de aprendizaje por refuerzo profundo desarrollados por OpenAI, que incluye tutoriales, guías y código de ejemplo para implementar y comprender los conceptos básicos del aprendizaje por refuerzo profundo.

- "Introduction to Reinforcement Learning" de Sutton and Barto: <http://incompleteideas.net/book/bookdraft2018jan1.pdf>

Este es un libro de texto clásico escrito por Richard Sutton y Andrew Barto, dos de los pioneros en el campo del aprendizaje por refuerzo. Proporciona una introducción completa y accesible a los conceptos y algoritmos del aprendizaje por refuerzo, con ejemplos prácticos y explicaciones claras.

2. Procesamiento de Datos en Tiempo Real: El procesamiento de datos en tiempo real es un tema avanzado que se refiere a la capacidad de analizar datos a medida que se generan, en lugar de procesarlos en lotes o de manera retrospectiva. Aprender sobre tecnologías y técnicas para procesar datos en tiempo real, como Apache Kafka, Apache Flink y sistemas de procesamiento de flujo, te permitirá trabajar en aplicaciones de alta velocidad como detección de fraudes en tiempo real, monitoreo de redes sociales y análisis de datos en tiempo real para la toma de decisiones empresariales.

- "Getting Started with Apache Kafka" de Apache Kafka documentation: <https://kafka.apache.org/quickstart>

Esta es la documentación oficial de Apache Kafka que proporciona una guía de inicio rápido para comenzar a trabajar con Kafka. Incluye una introducción a los conceptos básicos de Kafka, cómo configurarlo y cómo producir y consumir mensajes en tiempo real.

- "Introduction to Apache Flink" de Apache Flink documentation: [https://ci.apache.org/projects/flink/flink-docs-stable/docs/try-flink/local\\_installation/](https://ci.apache.org/projects/flink/flink-docs-stable/docs/try-flink/local_installation/)

Esta es la documentación oficial de Apache Flink que ofrece una introducción a Flink y cómo configurarlo en tu entorno local. Incluye una descripción general de los conceptos básicos de Flink, cómo instalarlo y cómo ejecutar aplicaciones de procesamiento de datos en tiempo real con Flink.

3. Aprendizaje Federado (Federated Learning): El aprendizaje federado es una técnica avanzada de aprendizaje colaborativo que permite entrenar modelos de aprendizaje en dispositivos distribuidos sin compartir los datos crudos. Aprender sobre cómo implementar y gestionar modelos de aprendizaje federado, incluyendo cuestiones de privacidad, seguridad y coordinación en entornos distribuidos, te permitirá trabajar en proyectos de aprendizaje en escenarios de colaboración descentralizada, como en la salud, la Internet de las cosas (IoT) y aplicaciones empresariales.

- "Federated Learning" de Google AI: <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>



Este es un artículo del blog de Google AI que presenta una introducción al aprendizaje federado. Cubre los conceptos básicos del aprendizaje federado, incluyendo la idea de entrenar modelos de aprendizaje en dispositivos distribuidos y colaborativos, y cómo se puede aplicar en casos de uso como el aprendizaje en dispositivos móviles.

- "Communication-Efficient Learning of Deep Networks from Decentralized Data" de Google Research:  
<https://arxiv.org/abs/1602.05629>

Este es un artículo de investigación de Google que describe el marco de aprendizaje federado y presenta un algoritmo de optimización para entrenar modelos de aprendizaje profundo en entornos federados. El artículo cubre los desafíos y soluciones del aprendizaje federado, incluyendo la comunicación eficiente entre dispositivos y la protección de la privacidad de los datos de los dispositivos participantes.

## 5. Conclusiones

En esta unidad hemos visto que la educación en la ciencia de datos es un proceso continuo de aprendizaje y adaptación a las nuevas tendencias y tecnologías en constante evolución. Una estrategia efectiva para seguir aprendiendo y destacarse en este campo es trabajar en proyectos prácticos y mostrarlos en GitHub, participar en competencias de Kaggle, mantenerse actualizado, desarrollar habilidades de comunicación y buscar oportunidades de empleo o proyectos freelance. Al aplicar estos consejos y estrategias, estarás mejor preparado para enfrentar los desafíos del mundo real de la ciencia de datos y alcanzar el éxito en tu carrera profesional.

Además, el próximo paso en el curso será el desarrollo del Trabajo Práctico Final Integrador, donde tendrás la oportunidad de aplicar todos los conocimientos vistos y consolidar tu comprensión de los conceptos y técnicas de la ciencia de datos. Te alentamos a revisar todo el contenido estudiado y embarcarte en este desafío, ya que te brindará una valiosa experiencia y te permitirá destacarte en tu formación como profesional en este campo en constante crecimiento. ¡Estamos seguros de que lo lograrás!