**Datos Generales**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **1. Nombre de la Asignatura**  Desarrollo de proyectos II | | **2. Nivel de formación**  Maestría | | **3. Clave de la Asignatura**  IH602 | |
| **4. Prerrequisitos**  Desarrollo de proyecto I | | **5. Área de Formación**  Especializante obligatoria | | **6. Departamento**  Sistemas de información | |
| **7. Modalidad**  Mixta | | **8. Tipo de Asignatura:**  Curso – Taller | | | |
| **9. Carga Horaria:** 60 horas | | | | | |
| **Teoría:** 48 horas | **Práctica:** 64 horas | | **Total:** 112 horas | | **Créditos:** 7 |
| **10. Trayectoria de la asignatura** | | | | | |

**Contenido del Programa**

|  |
| --- |
| **11 Presentación**  En este curso se presentarán y utilizarán de forma práctica herramientas básicas que sean útiles para desarrollar de manera efectiva proyectos relacionados con la ciencia de datos. |
| **12 Objetivos del programa**  **Objetivo General**  Conocer y utilizar herramientas prácticas para el desarrollo de proyectos de ciencia de datos. |
| **13 Contenido**  **Contenido temático**  1. Introducción a herramientas de Ciencia de Datos  2. Git y GitHub  3. Interfaz de usuario  4. Markdown  5. Análisis exploratorio de datos (EDA)  6. Scrape data en redes sociales  7. Procesamiento de lenguaje natural (NLP) básico  8. Spatial Data  9. Otros |
| ***Contenido desarrollado***  **1. Introducción a herramientas de Ciencia de Datos**  Objetivo particular: Descripción general de las herramientas que se revisarán  durante el curso.  **2. Git y GitHub**  Objetivo particular: Desarrollar proyectos individuales y en equipo utilizando el  sistema de control de versiones de forma local y remota  2.1. Inscripción a GitHub e instalación de Git  2.2. Básicos: Add, Commit, Push, Pull, Status  2.2. Clone  2.3. Fork  2.4. Pull requests  2.5. Branches  2.6. Conflictos  **3. Interfaz de usuario**  Objetivo particular: Guía rápida para uso la interfaz de usuario para navegar  en nuestros archivos directamente en el sistema operativo, y conectar  nuestro repositorios remotos en GitHub con repositorios locales en Git  **4. Markdown**  Objetivo particular: Guía rápida para uso de comandos en Markdown  **5. Análisis exploratorio de datos (EDA)**  Objetivo particular: Realizar análisis exploratorio de datos, el cual conlleva tener  un proyecto estructurado, limpiar los datos, tratamiento de *missing values*,  visualización y análisis de datos.  **6. *Scrape* *data* en redes sociales**  Objetivo particular: Revisión de librerías para consultar publicaciones, comentarios,  hashtags, etc., en redes sociales.  **7. Procesamiento de lenguaje natural (*NLP*) básico**  Objetivo particular: Análisis de texto a nivel de caracteres y tokens, revisando librerías en  Python como NLTK y spaCy, uso de regex, revisión de *n-grams*, análisis de sentimiento  y *wordclouds*.  **8. *Spatial Data***  Objetivo particular: Revisión de librerías usando datos georreferenciados para análisis  espacial.  **9. Otros**  Objetivo particular: Revisión de material relacionado con Ciencia de Datos proveniente  de distintas fuentes como papers, documentos de trabajo, entornos de desarrollo,  librerías de Python, preguntas en *Stack Overflow* y pláticas con invitados. |
| **14. Actividades prácticas**  Durante el semestre los alumnos realizarán presentaciones de diversos papers y librerías de Python. Además, en las clases se llevarán a cabo actividades en los alumnos participarán de forma activa. Al final de la clase, se realizará un breve quiz a manera de control de lectura. |
| **15 Metodología**  Actividades en clase, lecturas, presentaciones de alumnos, quizes y desarrollo de proyectos. |
| **16. Evaluación**  ✏️ **Evaluación general**   |  |  | | --- | --- | | 1) **Presentaciones** | 20% | | 2) **Quizes** | 20% | | 3) **Actividades** | 10% | | 4) **Proyecto de análisis EDA** | 20% | | 5) **Proyecto final** | 30% | | **Total** | **100%** |   ⭐️ **Puntos extras**   |  |  | | --- | --- | | A) **Proyecto personal de ciencia de datos**  *Desarrollo del proyecto en GitHub y divulgación en LinkedIn, redes sociales u otros canales.* | 5% | | B) **Publicación**  *Publicación de ‘Papers’ o de artículos en revistas de ciencia de datos como* [*Towards Data Science*](https://towardsdatascience.com/)*, divulgación de la ciencia, etc.* | 5% | | C) **Voluntariado**  Participación comprobable como voluntario en [*Omdena*](https://omdena.com/)*,* [*ONU*](https://app.unv.org/?type=online)*,* [*Data Science for Social Good*](https://www.solveforgood.org/)*, entre otros.* | 5% | | D) **Eventos**  *Participación comprobable en eventos relacionados con Ciencia de Datos como ‘Hackathon’, congresos y conferencias.* | 5% |   ⚠️ **Consideraciones de puntos extras:**   * *Se pueden participar más de una vez en cualquiera de los rubros de puntos extra (A, B, C y D), y por cada participación podrá ser acreedor hasta un 5% adicional sobre la evaluación final.* * *Máximos puntos extras a obtener:* ***20% adicional a la evaluación general****.* * *Para tener acceso a los puntos extras se necesita una calificación de al menos 75% en el “Total” de la Evaluación General.* |
| **17. Bibliografía**   * Bryan, J. “Happy Git and GitHub for the useR”. 2016. Link: https://happygitwithr.com * Katari, K. (Aug 21, 2020). Exploratory Data Analysis(EDA): Python. Towards Data Science. * UBC MDS public teaching materials. “Computing Platforms for Data Science”. 2022. Link: https://github.com/UBC-MDS/DSCI\_521\_platforms-dsci * Towards Data Science. 2022. Link: https://towardsdatascience.com * VanderPlas, J. “Python Data Science Handbook”. O'Reilly Media, 2016. * Link: https://jakevdp.github.io/PythonDataScienceHandbook/ * Wickham, H. and Grolemund, G. “R for Data Science”. O'Reilly Media, 2016. Link: https://r4ds.had.co.nz * Wilson, G., Bryan, J., Cranston, K., et al. (2016). "Good Enough Practices in Scientific Computing". Link: https://arxiv.org/abs/1609.00037   **Otros materiales**  ***Software y herramientas de apoyo***   * Python * Git & GitHub * Interfaz de usuario (e.g. PowerShell o GitBash en Windows, y Terminal en Mac) * Jupyter Lab, Jupyter Notebook, Visual Studio Code, RStudio |
| **19.- Perfil del profesor**  Profesor con nivel mínimo de maestría, con experiencia en el desarrollo de proyectos, programación y análisis de datos. |
| **20.- Nombre de los profesores que imparten la materia**  Víctor Hugo Cuspinera Contreras |
| **21.- Lugar y fecha de su aprobación (incluyendo la última actualización)** |
| **22.- Instancias que aprobaron el programa (Junta Académica y/ó Coordinación del programa)** |