	Computación	Docente: Diego Quisi Peralta
	SISTEMAS EXPERTOS – INTELIGENCIA ARTIFICIAL 1	Período Lectivo: Septiembre 2020 – Febrero 2021

		FORMATO DE GUÍA DE PRÁCTICA DE LABORATORIO / TALLERES / CENTROS DE SIMULACIÓN – PARA DOCENTES	
CARRERA: COMPUTACIÓN/INGENIERÍA DE SISTEMAS		ASIGNATURA: SISTEMAS EXPERTOS - INTELIGENCIA ARTIFICIAL 1	
NRO. PROYECTO:	1.1	TÍTULO PROYECTO: Proyecto Integrador Interciclo Desarrollo e implementación de un sistema de búsqueda, casos y/o similitud en una base de datos orientada a grafos.	
OBJETIVO: Reforzar los conocimientos adquiridos en clase sobre la búsqueda, casos y/o similitud para el diseño e implementación de sistemas inteligentes utilizando redes sociales en el contexto político del Ecuador.			
INSTRUCCIONES:		1. Revisar el contenido teórico del tema	
		2. Profundizar los conocimientos revisando los libros guías, los enlaces contenidos en los objetos de aprendizaje, cuadernos Python y la documentación disponible en fuentes académicas en línea.	
		3. Deberá desarrollar un sistema inteligente en base a redes sociales, por ejemplo twitter del presidente con mas seguidores, entre otros.	
		4. Deberá generar un informe empleando una herramienta Web 2.0 (Tutorial o manual técnico) .	
		5. Tomar en consideración que la evaluación del trabajo a realizarse de forma individual dependerá de los siguientes parámetros: Nivel de precisión y explicación de la propuesta planteada en cada uno de los algoritmos de búsqueda, similitud y casos (Cuadernos Jupyter). 50% Tutorial del sistema del uso de los algoritmos en Neo4j 25% (Pagina Web) Exposición 25% . 6. Fecha de presentación: 08 de Junio del 2021 a 23:55. 7. Puntos extras: Cualquier mejora, innovación o investigación adicional sera valorado como puntos extras directos al inter-ciclo.	
ACTIVIDADES POR DESARROLLAR			

1. Investigue, diseñe y desarrolle e implemente 1 y/o 2 algoritmos dentro de una base de datos orientadas a grafos del siguiente link (<https://neo4j.com/docs/graph-data-science/current/algorithms/>).

Problema: Se desea generar métodos inteligentes para encontrar información basada en un ejemplo aplicando datos del presidente obtenidos de redes sociales, etc..

Pasos a seguir:

1. Realizar una extracción de datos en base a palabras claves del presidente.
2. Migrar esta información a la base de grafos Neo4j, en virtud de ello, se deberá tener al menos 2000 nodos.
3. Implementar los algoritmos inteligentes dentro de la base de grafos.
4. Consultar información y/o resultados del grafo aplicando los algoritmos inteligentes.
6. Agregar conclusiones y recomendaciones.

Introducción:

Los algoritmos de grafos se utilizan para calcular métricas para grafos, nodos o relaciones.

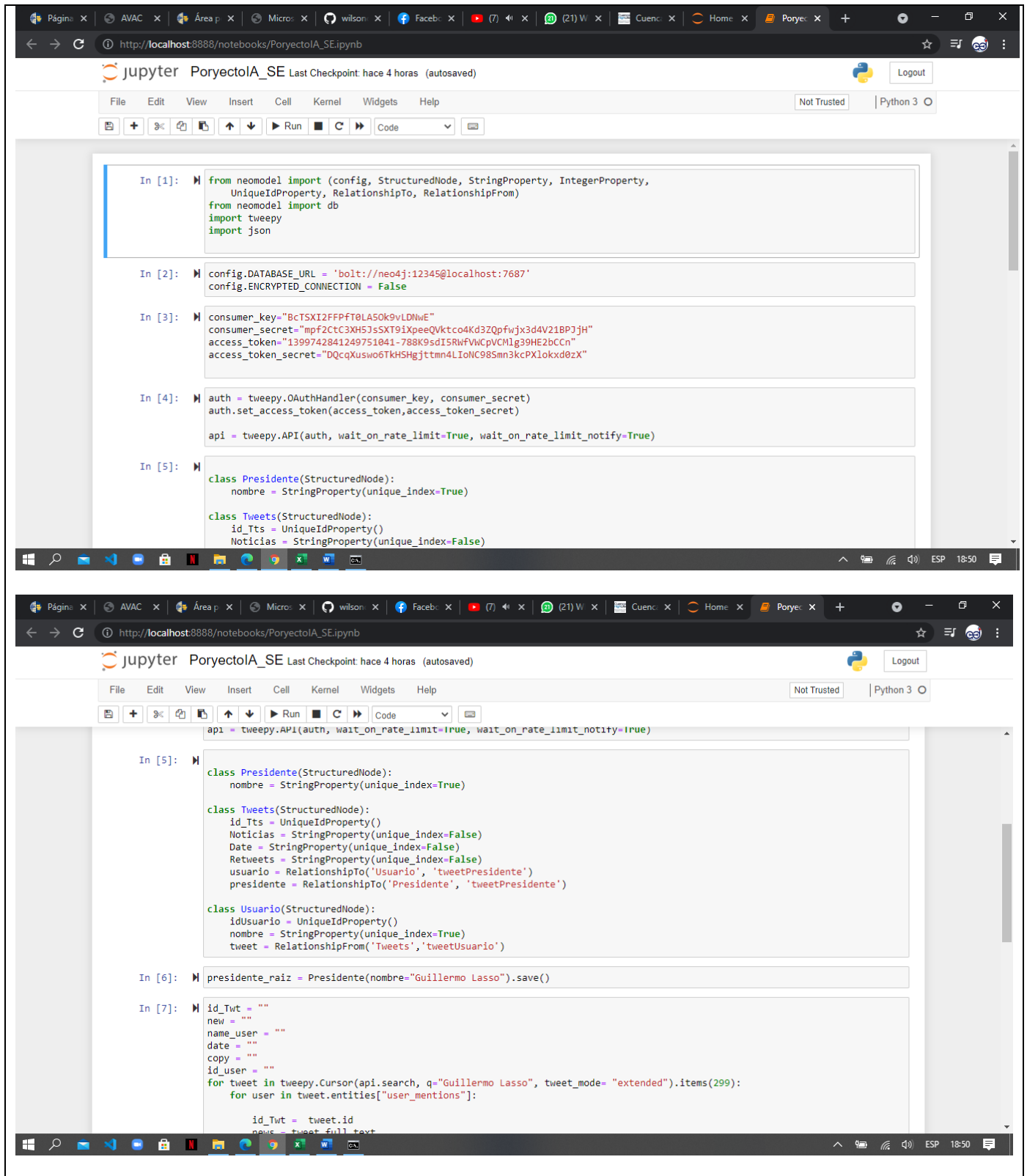
Pueden proporcionar información sobre entidades relevantes en el gráfico (centralidades, clasificación) o estructuras inherentes como las comunidades (detección de comunidad, partición de grafos, agrupación, etc.).

Muchos algoritmos de gráficos son enfoques iterativos que frecuentemente atraviesan el grafo para el cálculo utilizando caminatas aleatorias, búsquedas de amplitud o de profundidad, o coincidencia de patrones. En Neo4j Graph Data Science contiene una gran cantidad de algoritmos que se pueden aplicar sobre una base de datos orientadas a grafos:

- Sección 5.2, "Algoritmos de centralidad"
- Sección 5.3, "Algoritmos de detección de la comunidad"
- Sección 5.4, "Algoritmos de similitud"
- Sección 5.6, "Algoritmos de predicción de enlaces"

La **solución** se debe seleccionar 1 algoritmo (IA) y 2(SE) y generar un cuaderno que contenga la siguiente información:

- Introducción al tipo de algoritmo.
- Descripción del algoritmo.
- Ejemplificación usando datos reales (Redes Sociales).
- Resultados y análisis.
- Mejoras y recomendaciones(Material adicional o refuerzo del algoritmo).
- Conclusiones y trabajos futuros(Ejercicios/Áreas aplicables).



The image shows a Jupyter Notebook interface with the following code cells:

```
In [1]: from neomodel import (config, StructuredNode, StringProperty, IntegerProperty,
        UniqueIdProperty, RelationshipTo, RelationshipFrom)
        from neomodel import db
        import tweepy
        import json

In [2]: config.DATABASE_URL = 'bolt://neo4j:12345@localhost:7687'
        config.ENCRYPTED_CONNECTION = False

In [3]: consumer_key="BcTSXI2FFPFT0LA50k9vLDNwE"
        consumer_secret="mpf2CtC3XH5JsXT9iXpeeQVktco4Kd3ZQpfwjx3d4V218PJjH"
        access_token="1399742841249751041-788K9sdI5RwFVwCpVCM1g39HE2bCCn"
        access_token_secret="DQcqXuswo6TKH5Hgjttn4LIoNC985mn3kcPX1okxd0zX"

In [4]: auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
        auth.set_access_token(access_token, access_token_secret)
        api = tweepy.API(auth, wait_on_rate_limit=True, wait_on_rate_limit_notify=True)

In [5]: class Presidente(StructuredNode):
        nombre = StringProperty(unique_index=True)

        class Tweets(StructuredNode):
            id_Tts = UniqueIdProperty()
            Noticias = StringProperty(unique_index=False)

api = tweepy.API(auth, wait_on_rate_limit=True, wait_on_rate_limit_notify=True)

In [5]: class Presidente(StructuredNode):
        nombre = StringProperty(unique_index=True)

        class Tweets(StructuredNode):
            id_Tts = UniqueIdProperty()
            Noticias = StringProperty(unique_index=False)
            Date = StringProperty(unique_index=False)
            Retweets = StringProperty(unique_index=False)
            usuario = RelationshipTo('Usuario', 'tweetPresidente')
            presidente = RelationshipTo('Presidente', 'tweetPresidente')

        class Usuario(StructuredNode):
            idUsuario = UniqueIdProperty()
            nombre = StringProperty(unique_index=True)
            tweet = RelationshipFrom('Tweets', 'tweetUsuario')

In [6]: presidente_raiz = Presidente(nombre="Guillermo Lasso").save()

In [7]: id_Twt = ""
        new = ""
        name_user = ""
        date = ""
        copy = ""
        id_user = ""
        for tweet in tweepy.Cursor(api.search, q="Guillermo Lasso", tweet_mode="extended").items(299):
            for user in tweet.entities["user_mentions"]:
                id_Twt = tweet.id
                new = tweet.full_text
```

The screenshot shows a Jupyter Notebook running in a web browser at `http://localhost:8888/notebooks/PoryectoIA_SE.ipynb`. The notebook has a menu bar with File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. Below the menu is a toolbar with icons for file operations, code execution, and output viewing. The code cell is labeled 'In [7]:'. The Python code uses the Tweepy library to search for tweets mentioning 'Guillermo Lasso' and save the results into a database. The output of the code is displayed below the cell, showing the details of a specific tweet and its associated user.

```
In [7]: id_Twt = ""
new = ""
name_user = ""
date = ""
copy = ""
id_usuario = ""
for tweet in tweepy.Cursor(api.search, q="Guillermo Lasso", tweet_mode="extended").items(299):
    for user in tweet.entities["user_mentions"]:

        id_Twt = tweet.id
        news = tweet.full_text
        name_user = user["name"]
        id_usuario = user["id_str"]
        date = tweet.created_at
        copy = tweet.retweet_count
        print ("Registro: ")
        print ("Id:", id_Twt)
        print ("Noticias:", news)
        print ("Usuario:", name_user)
        print ("Fecha:", date)
        print ("Retweets:", copy)
        print ("IdUsuario:", id_usuario)
        print ("\n")

        tw = Tweets(id_TTS=id_Twt, Noticias=news, Date=date, Retweets=copy).save()
        us = Usuario(idUsuario=id_usuario, nombre=name_user).save()

        tw.usuario.connect(us)
        tw.presidente.connect(presidente_raiz)
```

Registro:
Id: 1403443451891437572
Noticias: RT @Expresoec: El presidente Guillermo Lasso asistió este 11 de junio de 2021 a su primera ceremonia militar tr
as su posesión el pasado 24..
Usuario: Diario Expreso
Fecha: 2021-06-11 20:06:08
Retweets: 2
IdUsuario 35047698

2. Tutorial técnico del uso y proceso de Neo4j (Manual técnico):

- Generar una pagina web:
 - Planteamiento y descripción del problema.
 - Proceso de solución.
 - Conclusiones y recomendaciones.
 - Resultados y validaciones.

PROYECTO INTERCICLO IA Y SE x +


← → C Archivo C:/Users/Usuario/Desktop/ProyectoIA_SE/index.html

ALGORITMOS

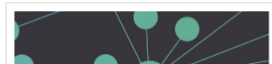
Community Detection AlgorithmsS Centrality Algorithms

Community Detection Algorithms

Los algoritmos de detección de comunidades se utilizan para evaluar cómo los grupos de nodos se agrupan o se dividen, así como su tendencia a fortalecerse o separarse.



Centrality Algorithms




PROYECTO INTERCICLO IA Y SE x +

← → C Archivo C:/Users/Usuario/Desktop/ProyectoIA_SE/index.html

Community Detection Algorithms

Los algoritmos de detección de comunidades se utilizan para evaluar cómo los grupos de nodos se agrupan o se dividen, así como su tendencia a fortalecerse o separarse.



Centrality Algorithms

Los algoritmos de centralidad se utilizan para determinar la importancia de distintos nodos en una red.

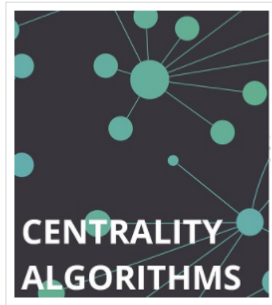
La biblioteca Neo4j GDS incluye los siguientes algoritmos de centralidad, agrupados por nivel de calidad:

CALIDAD DE PRODUCCION

- Rango de página (articleRank)
- Clasificación del artículo
- Centralidad del vector propio
- Centralidad de intermediación
- Centralidad de grado

ALFA

- Cercanía Centralidad
- Centralidad armónica



PROYECTO INTERCICLO IA Y SE x +

← → C Archivo C:/Users/Usuario/Desktop/ProyectoIA_SE/index.html

PROYECTO INTERCICLO IA Y SE

Planteamiento y descripción del problema

PROBLEMA 1
Se trabajara con el primer algoritmo: Community Detection Algorithms

PROBLEMA 2
Se trabajara con el segundo algoritmo: Centrality Algorithms(articleRank)

Problema 1 y Solucion

Community Detection Algorithms(Louvain)

Aqui buscaremos todos los tweets del presidente y sus respectivos usuarios, para ellos plantearemos de la siguiente forma la solucion.

Primero:

Configuramos Nuestro algoritmo

Segundo:

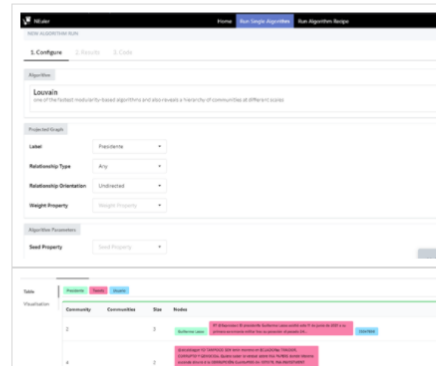
Despues de la configuracion nos dara ya un resultado

Tercero:

El algoritmo nos lanzara un codigo el cual debemos correrlo en nuestro neo4j

Cuarto:

Corremos el algoritmo en neo4j y nos dara los siguientes resultados incluido



PROYECTO INTERCICLO IA Y SE

file:///C:/Users/Usuario/Desktop/ProyectoIA_SE/index.html

Problema 2 y Solucion

Centrality Algorithms(article rank)

Aqui buscaremos todos los tweets del presidente y sus respectivos usuarios, para ellos plantearemos de la siguiente forma la solucion.

Primero:

Configuramos Nuestro algoritmo

Segundo:

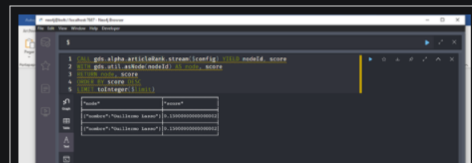
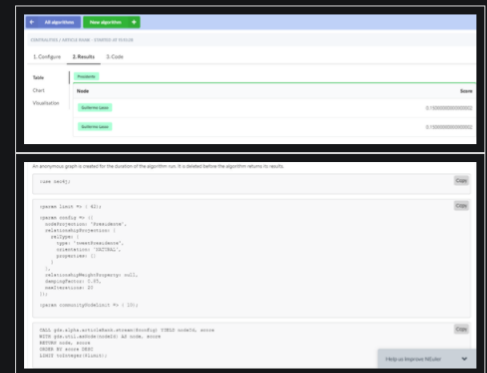
Despues de la configuracion nos dara ya un resultado

Tercero:

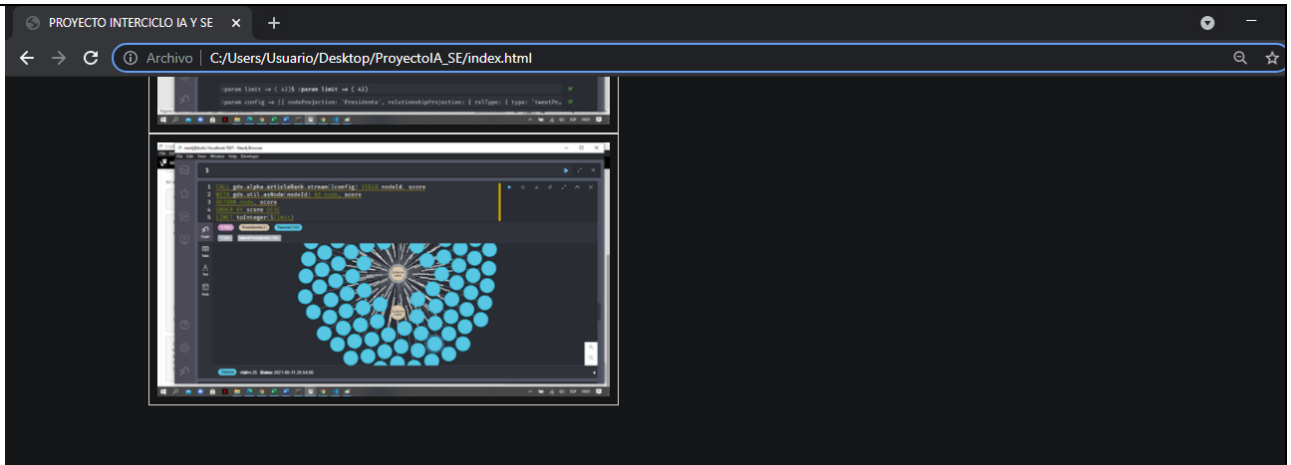
El algoritmo nos lanzara un codigo el cual debemos correrlo en nuestro neo4j

Cuarto:

Corremos el algoritmo en neo4j y nos dara los siguientes resultados incluido el grafico.



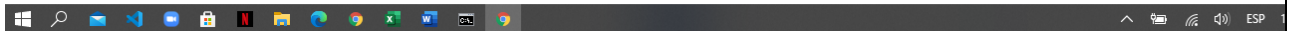
PROYECTO INTERCICLO IA Y SE



Conclusiones y Recomendaciones

El uso del comando neo4j-import solo se puede usar en la primera carga con base de datos vacía no es posible su uso para cargar incrementales. La ventaja es la diversidad de ficheros tanto nodos como relaciones que se pueden cargar de una vez. La desventaja es imposibilidad de crear índices.

Las ventajas de uso de neo4j-import es la gestión de errores ya que un error en la carga de un nodo no supone la cancelación de toda la carga y la información de ellos en el fichero bad.log que se crea en la carpeta de la base de datos.



RESULTADO(S) OBTENIDO(S):

- Interpreta de forma correcta los métodos de búsqueda, similitud, centralidad y predicción de enlaces y su aplicabilidad.
- Identifica correctamente qué herramientas de búsqueda orientadas a grafos se pueden aplicar.

CONCLUSIONES:

- Los estudiantes identifican las principales estructuras para la búsqueda, similitud, centralidad y predicción.
- Los estudiantes implementan soluciones de soporte a la toma de decisiones basadas en grafos.

RECOMENDACIONES:


- Revisar la información proporcionada por el docente previo a la práctica.
- Haber asistido a las sesiones de clase.
- Consultar con el docente las dudas que puedan surgir al momento de realizar la práctica.


BIBLIOGRAFIA:

- [1]: <https://neo4j.com/docs/graph-data-science/current/algorithms/>
- [2]: <https://neo4j.com/graph-data-science-library/>
- [3]: <https://neo4j.com/blog/story-behind-russian-twitter-trolls/>
- [4]: <https://www.freecodecamp.org/news/monitoring-the-french-presidential-election-on-twitter-with-python-6a2a9310e6f4/>
- [5]: <https://lib.dr.iastate.edu/etd/15949/>

Docente / Técnico Docente: Ing. Diego Quisi Peralta Msc.

Firma: _____

	Computación	Docente: Diego Quisi Peralta
	SISTEMAS EXPERTOS – INTELIGENCIA ARTIFICIAL 1	Período Lectivo: Septiembre 2020 – Febrero 2021

		FORMATO DE INFORME DE PRÁCTICA DE LABORATORIO / TALLERES / CENTROS DE SIMULACIÓN – PARA ESTUDIANTES	
CARRERA:		ASIGNATURA:	
NRO. PRÁCTICA:		TÍTULO PRÁCTICA:	
OBJETIVO ALCANZADO:			
<p style="text-align: center;">ACTIVIDADES DESARROLLADAS</p>			
1.			
2.			
3.			
4.			
5.			
6.			
N.			
RESULTADO(S) OBTENIDO(S):			
CONCLUSIONES:			
RECOMENDACIONES:			

Nombre de estudiante: _____

Firma de estudiante: _____