



UNIVERSIDAD  
CENTRAL

LIBROS

Referencias sugeridas

---

Jesús Alfredo Sanabria Mejía

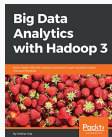
0060 Bases de Datos

17 de septiembre de 2022

# Libros Observados

1. Big Data Analytics with Hadoop 3.
2. Patterns identification and data mining in weather and climate
3. Hands-On Python Natural Language Processing
4. Next generation databases: nosql and big data

# Big Data Analytics with Hadoop 3



[Alla, 2018]

- Editorial Packt Publishing (472 paginas)
- Edición (29/05/2018)
- Idioma Inglés
- ISBN 781788628846



Sridhar Alla: es un Licenciado en Informática Indio, experto en big data que ayuda a las empresas a resolver problemas complejos en sistemas de computo distribuidos, ciencia de datos a gran escala y práctica de análisis.

El libro es una guía práctica de Apache Hadoop para el procesamiento de big data y enseña a combinar con una gran cantidad de otras herramientas de big data para crear soluciones de análisis. cuenta con la orientación mediante una ayuda de ejemplos prácticos. con códigos gratuitos que se puede descargar desde github

<https://github.com/packtpublishing/big-data-analytics-with-hadoop-3>

# Big Data Analytics with Hadoop 3



- 1. Introducción a Hadoop
- 2. Descripción general de Big Data Analytics
- 3. Procesamiento de Big Data con Mapreduce
- 4. Computación Científica y Grandes Análisis de datos con Python y Hadoop
- 5. Informática estadística de grandes datos con R y Hadoop
- 6. Análisis por lotes con Apache Spark
- 7. Análisis en tiempo real con Apache Spark
- 8. Análisis por lotes con Apache Flink
- 9. Procesamiento de Stream con Apache Flink
- 10. Visualización de Big Data (Herramientas)
- 11. Introducción al Computación en la Nube
- 12. Utilización de Servicios Web en Amazon.

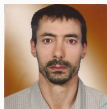
# Patterns identification and data mining in weather and climate



[Hannachi, 2021]

- Editorial Springer (625 paginas)
- Edición (2021)
- Idioma Inglés
- ISBN 9783030670726

El libro presenta diferentes perspectivas disponibles, para analizar series temporales multivariadas en la ciencia del clima para identificar patrones de variabilidad. El libro analiza diferentes métodos para identificar patrones de campos espaciotemporales, es en suma muy teorico, aunque cuenta con unos apartados a codigos en matlab de analisis, y referencias a sitios externos para investigar mas del tema.



Abdelwaheb Ben Ahmed

Hannachi: es un profesor asociado en la universidad de estocolmo experto en temas de climatologia estadistica, pronostico y asimilación de datos, estudios climaticos entre otros.

# Patterns identification and data mining in weather and climate (17 capitulos)



- 1.2 Exploración de datos, minería de datos y extracción de características
- 1.3 Principales preocupaciones en el análisis de datos climáticos
- 17. Machine Learning
- 17.2 Redes Neuronales
- A. Técnicas de suavizado.
- H. Enlaces a material y recursos de software

En general, es un libro demasiado teórico, con mucha formulación de modelos matemáticos y justificación de las técnicas para la extracción de información sobre patrones climáticos.

# Hands-On Python Natural Language Processing



[Kedia and Rasu, 2020]

- Editorial BIRMINGHAM - MUMBAI (304 paginas)
- Edición (2020)
- Idioma Inglés
- ISBN 978-1-83898-959-0



Mayank Rasu Vicepresidente, Análisis Cuantitativo - Desarrollo de Modelos en Barclays Investment Bank tiene más de 12 años de experiencia global como científico de datos y analista cuantitativo en la industria de banca de inversión.

Aman Kedia es promotor de la Inteligencia artificial (IA) y los algoritmos que la soportan. Ha trabajado en problemas del procesamiento del lenguaje natural (NLP), que abarcan la comparación de currículums y asistentes digitales, entre otros. Ha trabajado en Oracle y SAP, tratando de resolver problemas que aprovechan los avances en IA.

Este libro presentará información del Procesamiento de Lenguaje Natural PNL y sus aplicaciones, junto con modernas bibliotecas de Python que sirven de ayuda a ejemplos prácticos, para crear aplicaciones de NLP. El libro cubre varias metodologías y desafíos en la implementación de aplicaciones de NLP. Cubre tareas clave de NLP, como la clasificación de texto, la incrustación semántica, el análisis de sentimientos, la traducción automática y el desarrollo de un chatbot utilizando técnicas de aprendizaje automático y aprendizaje profundo.

# Hands-On Python

## Natural Language Processing (11 capítulos)

Los códigos de Ejemplo son de acceso libre por github

<https://github.com/packtpublishing/hands-on-python-natural-language-processing>

- 1. conceptos básicos de PNL
- 2. PNL usando Python
- 3. Construyendo vocabulario de PNL
- 4. Transformar texto en Estructuras de datos
- 5. incrustaciones de palabras y Mediciones para Textos
- 6. Exploración de sentencias en documentos e incrustaciones a nivel de documento y carácter.
- 7. Identificación de patrones en textos usando machine learnig
- 8. De las neuronas humanas a Neuronas artificiales para Comprender el texto
- 9. Aplicando Convoluciones l texto
- 10. Capturando relaciones temporales en textos
- 11. Estado del Arte en NLP



# Next generation databases: nosql and big data



[Harrison, 2020]

- Editorial APRESS / Springer (381 paginas)
- Edición (2020)
- Idioma Inglés
- ISBN 978-1-4842-1330-8



Guy Harrison (Australia) comenzó a trabajar como desarrollador de bases de datos a mediados de la década de 1980. Él ha escrito numerosos libros sobre desarrollo y rendimiento de bases de datos. Se unió a Quest Software (ahora parte de Dell) en 2000 y actualmente lidera el equipo que desarrolla Toad, Spotlight y Shareplex familias de productos.

Este libro describe para qué se diseñó cada tecnología de Base de Datos. Muestra cómo se puede usar cada tecnología para resolver problemas reales de aplicaciones y negocios. Lo que es más importante, este libro destaca las diferencias arquitectónicas entre las tecnologías que son los factores críticos a considerar al elegir una plataforma de base de datos para proyectos nuevos y futuros.

# Next generation databases: nosql and big data (13 capítulos)



UNIVERSIDAD  
CENTRAL

- Bases de datos de próxima generación
- Tres revoluciones de bases de datos
- Google, grandes datos y Hadoop
- Sharding(fragmentación), Amazon y el nacimiento de NoSQL
- Bases de datos de documentos
- Las tablas no son tus amigas: bases de datos de gráficos
- Bases de datos de columnas
- ¿El fin del disco? SSD y bases de datos en memoria
- Patrones de bases de datos distribuidas
- Modelos de consistencia
- Modelos de datos y almacenamiento
- Lenguajes e Interfaces de Programación

# Recomendations



Alla, S. (2018).

*Big Data Analytics with Hadoop 3: Build highly effective analytics solutions to gain valuable insight into your big data.*

Packt Publishing, Birmingham, England.



Hannachi, A. (2021).

*Patterns identification and data mining in weather and climate.*

Springer Nature, Cham, Switzerland, 1 edition.



Harrison, G. (2020).

*Next generation databases: Nosql, NewSQL, and Big Data.*

Apress.



Kedia, A. and Rasu, M. (2020).

*Hands-On Python Natural Language Processing: Explore tools and techniques to analyze and process text with a view to building real-world NLP applications.*

Packt Publishing, Birmingham, England.

Thanks!