



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

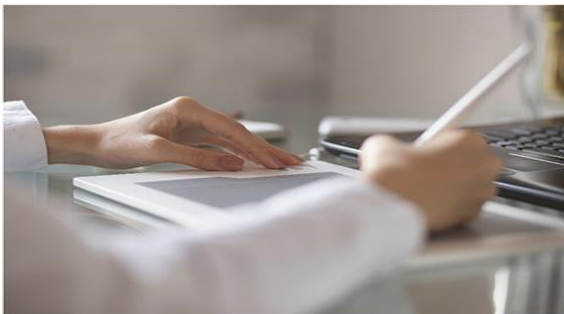
EDUCACIÓN
PROFESIONAL

Diplomado Big Data y Analítica de Datos 2021

Curso: Fundamentos Machine & Deep Learning

Educación Profesional Escuela de Ingeniería UC

Profesor: Rodrigo Sandoval



OBJETIVOS

Entender en qué consiste el Aprendizaje Automático o Machine Learning (ML) identificando potencias y limitaciones para resolver diferentes tipos de problemas en diferentes industrias.

Comprender y visualizar situaciones y su contexto en las cuáles ML podrá resolver problemas de gran valor de negocio.

Ser capaz de desarrollar un plan de implementación de proyectos y/o capacidades de ML en una empresa, analizando y evaluando infraestructura, servicios, y disponibilidad de datos.

Entender y comprender el aporte de Deep Learning (y sus ventajas sobre Machine Learning).

Visualizar cómo estas tecnologías pueden apoyar un plan estratégico de desarrollo de innovación basada en ML, alineado con la estrategia corporativa.

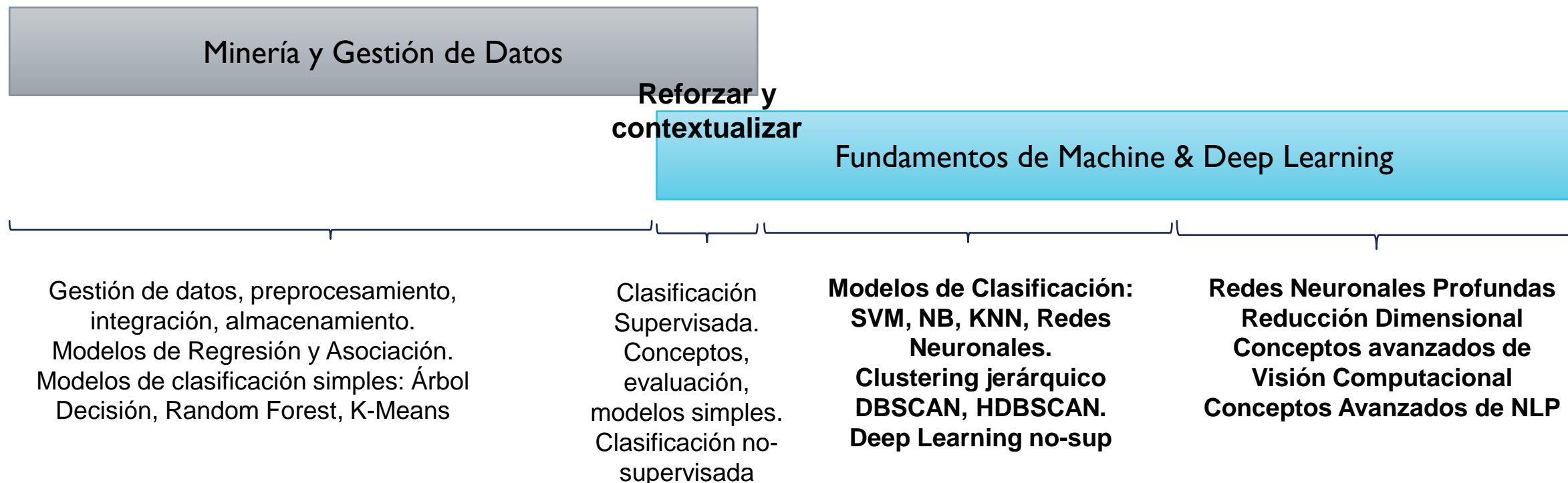
MACHINE LEARNING

- Introducción, Motivación, Estado Actual
 - Conceptos. Big Data. IA. Proceso desarrollo y modelos.
- Machine Learning Supervisado
 - Métodos de Entrenamiento y Evaluación.
 - Modelos: KNN, Naive Bayes, SVM, Redes Neuronales.
 - Control de overfitting.
 - Regresión no-lineal y logística.
- Machine Learning No-Supervisado
 - DBSCAN, HDBSCAN.
- Proyectos Machine/Deep Learning

DEEP LEARNING

- Deep Learning Supervisado
 - Características de las Redes Neuronales Profundas. Intro a Álgebra Lineal.
 - Funciones activación, Backpropagation.
 - Tipos de Redes Profundas.
- Visión Computacional con Redes Profundas
 - Redes Convolucionales
 - Reducción Dimensional
- Procesamiento de Lenguaje Natural
 - Vectorización de texto.
 - Secuenciamiento y Atención. Transformers.
- Deep Learning No-Supervisado
 - Autoencoders, RBMs, GAN.

RELACIÓN ENTRE CURSOS



Ambos cursos son complementarios, aunque seguir la secuencia normal es una gran ayuda a entender mejor el contexto integral y avanzado de la inteligencia de datos.

METODOLOGÍA

- Clases expositivas.
 - Material en PDF de las presentaciones de clases.
- Actividades prácticas en computador para aplicar las técnicas aprendidas en datos reales.
 - 5 ejercicios en grupo, aplicando conceptos en diferentes sesiones
- Lecturas complementarias.
- Controles de conocimientos.
 - Un control y un examen al final del curso.

EVALUACIÓN

La evaluación de los aprendizajes desarrollados en el curso se realizará mediante:

- 5 ejercicios prácticos en computador, con objetivos a cumplir. 4 de ellos evaluados con nota. Se elimina el peor de todos y el promedio de los que quedan es un 50% de la nota final.
- Un control de conocimientos al finalizar sesión 4, con el 20% de la nota final.
- Un examen final, que pondera 30% de la nota final, que se realiza en la última sesión.



DINÁMICA EN CLASES

Clases y Ejercicios

Se presentan los conceptos y materia, fomentando la discusión por medio de preguntas y casos de ejemplo

¡Participen!

Hagan preguntas. Levanten la mano. Activen el **cámara** y **micrófono** y comenten.

Cámara encendida (mientras se pueda)

Pueden mantener su cámara apagada si tienen limitaciones de ancho de banda.

Clases Grabadas

Se graban todas las clases, pero el mayor valor es asistir en tiempo real y participar en la discusión y ejercicios

Pausas

Tendremos dos pausas de 10 minutos durante la sesión de clases.

CALENDARIO GENERAL

#	Contenido	Actividades
1	Presentación. Introducción, motivación y conceptos esenciales.	(Lectura 1, fuera de clases)
2	Regresión No-Lineal y Logística. Clasificación Supervisada. Métodos de evaluación. Modelos y Overfitting.	Ejercicio 1: Clasificación supervisada y aplicación mejoras
3	Clasificación Supervisada, 2ª parte. Otros modelos. Redes Neuronales, introducción al Deep Learning.	Ejercicio 2: Configuración Red Neuronal multi-capa
4	Visión Computacional. Reducción Dimensional.	Control
5	Introducción al Procesamiento de Lenguaje Natural (NLP)	Ejercicio 3: Clasificación Texto
6	Modelos avanzados de NLU y conceptos de secuencia en redes	Ejercicio 4: Generación Texto
7	Modelos no-supervisados.	Ejercicio 5: Clustering
8	Proyectos de Machine Learning y estado del arte.	Examen Final

BIBLIOGRAFÍA

Efficient Learning Machines

Theories, Concepts, and Applications for Engineers and System Designers

Efficient Learning Machines

- Theories, Concepts and Applications for Engineers and System Designers
- M.Awad & R. Khanna
- 2015, Apress Open

Evaluating Machine Learning Models

A Beginner's Guide to Key Concepts and Pitfalls

Evaluating Machine Learning Models

- A Beginner's Guide to Key Concepts and Pitfalls
- Alice Zheng
- 2015, O'Reilly Open

UNDERSTANDING MACHINE LEARNING

FROM THEORY TO ALGORITHMS

Understanding Machine Learning

- S. Shalev-Shwartz & S. Ben-David
- 2014, Cambridge University Press



Machine Learning and Big Data

Intuitive ML and Big Data in C++, Scala, Java and Python
(Online)
2021, Kareem Alkaseer

Artificial Neural Networks

B. MEHLIG

Artificial Neural Networks

- Lecture Notes
- Bernard Mehlig
- 2019, Göteborg, Sweden

MACHINE LEARNING YEARNING

Technical Strategy for AI Engineers, In the Era of Deep Learning

Machine Learning Yearning

- Andrew Ng
- 2019, deeplearning.ai

Acceso a la Bibliografía

- Efficient Learning Machines.
<https://link.springer.com/book/10.1007/978-1-4302-5990-9>
- Evaluating Machine Learning Models
<https://www.oreilly.com/library/view/evaluating-machine-learning/9781492048756/>
- Understanding Machine Learning
<https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/>
- Machine Learning and Big Data. Alkaseer.
<http://www.kareemalkaseer.com/books/ml/>
- Artificial Neural Networks
<https://arxiv.org/abs/1901.05639> (<https://arxiv.org/pdf/1901.05639.pdf>)
- Machine Learning Yearning
<https://www.deeplearning.ai/machine-learning-yearning/>

Ingeniero Civil de Industrias, PUC
mención Ciencia de la Computación (1995)

Master of Science
Inteligencia Artificial (1996)

Desde 1991
Desarrollo software
Gestión proyectos innovación

Desde 1996
Profesor asociado adjunto
Depto. Ciencia Computación
Ingeniería, PUC

2006-2011
SYNOPSYS INC.
Desarrollo IA para
industria de los microchips



RodrigoSandoval.net



rsandova@ing.puc.cl

rodrigo@RSolver.com



@RSandovalSolver



/in/RodrigoSandoval

Desde 2011 **R:Solver**
Desarrollo de IA en diferentes industrias.
Papers y patente

Using a Genetic Algorithm for
exploring a Semiconductor Fabrication Model

Rodrigo Sandoval
TCAD Technical Lead
Synopsys R&D Center
Santiago, Chile
rodrigo.sandoval@synopsys.com

Abstract—Initial iterations or generations of this particular Genetic Algorithm (GA) implementation for optimizing a semiconductor fabrication process, not only search for an optimal solution, but also, with the aid of a visual tool, help

- A user-defined weight is assigned to each response, representing its relative importance in the goal function result.
- These normalized and desirability-adapted values are then used to calculate the final goal function result, which is then used to determine the next iteration.

setup of an optimization model in the following semiconductor process, the model will be used for searching deposition, using High-Density Model [2].

on of a deposition process on understanding and ultimately for a specific semiconductor operating a given deposition

deposition, used by machines of calculated at every point on the three-dimensional integrations of a this case, the distribution is is the parameter exponent.

flux becomes a source of the city proportional to the ion flux, angle between the surface normal size of ions.

removed from surface with redeposition is regular cosine distribution.

flux becomes a source of the city proportional to the ion flux, angle between the surface normal size of ions.

Approved for use through 02/21/2014. OMB 0655-0032
U.S. Patent and Trademark Office U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. _____

INVENTOR(S)		
Given Name (first and middle if any)	Family Name or Surname	Residence (City and either State or Foreign Country)
Rodrigo Andrés	SANDOVAL URRICH	Santiago, Chile
Juan Ignacio	SAA HARGOUS	Santiago, Chile

Additional inventors are being named on the _____ separately numbered sheets attached hereto.

TITLE OF THE INVENTION (500 characters max):
DOCUMENT CHARACTERIZATION METHOD

Direct all correspondence to: _____

☒ The address corresponding to Customer Number: **30593**

OR

☐ Firm or Individual Name _____

Address _____

City _____ State _____ Zip _____

Country _____ Telephone _____ Email _____

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Application Data Sheet. See 37 CFR 1.76. ☐ CD(s), Number of CDs _____

☐ Drawings(s). Number of Sheets _____ ☐ Other (specify) _____

☒ Specification (e.g., description of the invention). Number of Pages _____

Fees Due: Filing fee of \$500 (\$100 for small entity) (See 37 CFR 1.53(c)). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is \$400 (\$200 for small entity) (\$100 for micro entity) for each additional 50 sheets or fraction thereof (see 35 U.S.C. 431(d)(1)(G) and 37 CFR 1.361).

METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant certifies small entity status. See 37 CFR 1.27.

☐ Applicant must attach form PTO/SB-15A or B or equivalent.

☐ A check or money order made payable to the Director of the United States Patent and Trademark Office is enclosed to cover the filing fee and application size fee (if applicable).

☐ Payment by credit card, form PTO-2558 is attached.

☒ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: 06-0790

TOTAL FEE AMOUNT (\$) **\$130.00**

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.53. The information is required to obtain or retain a benefit in the patent which is to be filed by the USPTO to present an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.121 and 1.14. This collection is estimated to take 10 hours to complete, including gathering, reviewing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1480, Alexandria, VA 22303-1480. DO NOT SEND PTO'S OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1480, Alexandria, VA 22303-1480.

If you need assistance in completing the form, call 1-800-PTO-9196 and select option 2.

MASTER OF SCIENCE 1995-1996

“... OPTIMIZACIÓN
BASADA EN TECNICAS
NO CONVENCIONALES
DE INTELIGENCIA
ARTIFICIAL”

FONDOS FONDECYT



DESARROLLO DE UNA HERRAMIENTA DE OPTIMIZACION BASADA EN TECNICAS NO CONVENCIONALES DE INTELIGENCIA ARTIFICIAL, INVESTIGACION OPERATIVA Y COMPUTACION PARALELA: SEGUNDA PARTE

Nº Proyecto : 1950920	Concurso : FONDECYT-REGULAR
Año Concurso : 1995	Consejo : TECNOLOGIA
Duración : 2 años	Estado : APROBADO
Inicio : Marzo 1995	Término : Marzo 1997
Recursos Asignados (en miles de pesos del año de concurso)	Año 1995: 13.879 Año 1996: 13.488

INVESTIGADOR RESPONSABLE

- NUSSBAUM VOEHL, MIGUEL

COLABORADOR(ES)

- CAMPOS ULLOA, ALVARO ENRIQUE
- WEINTRAUB POHORILLE, ANDRES FELIX

TESISTA(S) ASOCIADO(S) AL PROYECTO

- BUZETA ARAYA, RODRIGO JAVIER
- FISCHER GESCHE, ROBERT MICHAEL
- SANDOVAL URRICH, RODRIGO ANDRES

INSTITUCION PRINCIPAL

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE (PUC)
FACULTAD DE INGENIERIA
ESCUELA DE INGENIERIA
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACION

GRACIAS



rsandova@ing.puc.cl

rodrigo@RSolver.com



@RSandovalSolver



/in/RodrigoSandoval

www.RodrigoSandoval.net

www.RSolver.com