

# Postgraduate Diploma

Machine Learning and Data Mining  
Techniques in Genomic Oncology





## Postgraduate Diploma

### Machine Learning and Data Mining Techniques in Genomic Oncology

Course Modality: **Online**

Duration: **6 months**

Certificate: **TECH Technological University**

Official N° of hours: **600 h.**

Website: [www.techtute.com/pk/medicine/postgraduate-diploma/postgraduate-diploma-machine-learning-dara-mining-techniques-genomic-oncology](http://www.techtute.com/pk/medicine/postgraduate-diploma/postgraduate-diploma-machine-learning-dara-mining-techniques-genomic-oncology)

# Index

01

Introduction

---

*p. 4*

02

Objectives

---

*p. 8*

03

Course Management

---

*p. 12*

04

Structure and Content

---

*p. 20*

05

Methodology

---

*p. 24*

06

Certificate

---

*p. 32*

# 01

# Introduction

The concept of genomic or precision oncology is not entirely new; physicians have been using blood type to tailor blood transfusions for over a century. What is different today is the rapid growth of genomic data that can be collected quickly and cheaply from the patient and the community at large, and the potential to gain insights from sharing that data.



“

*This program will provide you with a sense of confidence in medical practice, which will help you grow personally and professionally”*

The scale and complexity of genomic data dwarf the measurements traditionally used in laboratory testing. In recent years there has been an enormous development of informatics to analyze and interpret DNA sequencing, and it has created a gap between biological knowledge and its application to routine clinical practice. It is therefore necessary to educate, disseminate and incorporate these informatics techniques among the medical community in order to be able to interpret the massive analysis of data from publications, biological or medical databases and medical records, among others, and thus enrich the biological information available at the clinical level.

This machine learning will enable the development of precision oncology, in order to interpret genomic characteristics and find targeted therapies, or to identify risks to certain diseases and establish more individualized preventive measures. A fundamental objective of the program is to bring students closer to and disseminate computer knowledge, which is already applied in other fields of knowledge, but has minimal implementation in the medical world, despite the fact that for genomic medicine to become a reality, it is necessary to accurately interpret the huge volume of clinical information currently available and associate it with the biological data generated after a bioinformatic analysis. While this is a difficult challenge, it will allow the effects of genetic variation and potential therapies to be explored quickly, inexpensively and with greater precision than is currently possible.

Humans are not naturally equipped to perceive and interpret genomic sequences, to understand all the mechanisms, pathways and interactions that take place within a living cell, nor to make medical decisions with tens or hundreds of variables. To move forward, a system with superhuman analytical capabilities is required to simplify the work environment and show the relationships and proximities between variables. In genomics and biology, it is now recognized that it is better to spend resources on new computational techniques than on pure data collection, something that is possibly the same in medicine and, of course, oncology.

This **Postgraduate Diploma in Machine Learning and Data Mining Techniques in Genomic Oncology** contains the most complete and up-to-date scientific program on the market. The most important features include:

- ♦ Development of case studies presented by experts in Machine Learning and Data Mining Techniques in Genomic Oncology
- ♦ Its graphic, schematic and practical contents are designed to provide scientific and practical information on those disciplines that are essential for professional practice
- ♦ News on Machine Learning and Data Mining Techniques in Genomic Oncology
- ♦ It contains practical exercises where the self-assessment process can be carried out to improve learning
- ♦ With special emphasis on innovative methodologies in Machine Learning and Data Mining Techniques in Genomic Oncology
- ♦ All of this will be complemented by theoretical lessons, questions to the expert, debate forums on controversial topics, and individual reflection assignments
- ♦ Content that is accessible from any fixed or portable device with an Internet connection



*Update your knowledge with the Postgraduate Diploma Program in Machine Learning and Data Mining Techniques in Genomic Oncology"*

“

*This Postgraduate Diploma may be the best investment you can make in the selection of a refresher program for two reasons: in addition to updating your knowledge in Machine Learning and Data Mining Techniques in Genomic Oncology, you will obtain a qualification from TECH Technological University"*

Its teaching staff includes professionals belonging to the field of Machine Learning and Data Mining Techniques in Genomic Oncology, who bring to this program the experience of their work, as well as recognized specialists belonging to reference societies and prestigious universities.

The multimedia content developed with the latest educational technology will provide the professional with situated and contextual learning, i.e., a simulated environment that will provide an immersive program to learn in real situations.

This program is designed around Problem-Based Learning, whereby the student must try to solve the different professional practice situations that arise during the Programming. For this, the student will be assisted by a novel interactive video system developed by recognized experts in the field of Machine Learning Techniques and Data Mining in Genomic Oncology with extensive teaching experience.

*Increase your decision-making confidence by updating your knowledge with this Postgraduate Diploma.*

*Take the opportunity to learn about the latest advances in Machine Learning and Data Mining Techniques in Genomic Oncology and improve the care of your patients.*



# 02 Objectives

The program in Machine Learning and Data Mining Techniques in Genomic Oncology is aimed at facilitating the performance of the physician dedicated to the treatment of oncological pathology, in which it is necessary to accurately interpret the huge volume of clinical information currently available and associate it with the biological data generated after a bioinformatics analysis.





A Newton's cradle with several silver spheres suspended by thin rods, set against a blue background with a grid pattern. The cradle is positioned in the upper left and center of the page, with the spheres appearing to be in motion.

“

*This program is designed for you to update your knowledge in Machine Learning and Data Mining Techniques in Genomic Oncology, with the use of the latest educational technology, to contribute with quality and confidence to decision-making"*



## General Objective

---

- ♦ Be able to accurately interpret the volume of clinical information currently available and associated with the biological data generated after a bioinformatic analysis



*Take the opportunity and take the step to get up to date on the latest developments in Machine Learning and Data Mining Techniques in Genomic Oncology"*





## Specific Objectives

---

### Module 1. Molecular Biology

- ♦ Update knowledge on the molecular biology of cancer, in relation to different concepts such as genetic heterogeneity or microenvironment reprogramming
- ♦ Understand what machine learning is and to use some of the techniques for data classification (decision tree, k-NN, Support Vector Machines, neural networks, etc.)
- ♦ Learn how to divide data into a test set and a training set, and discover the concepts of bias and variance

### Module 2. Data Mining Applied to Genomics

- ♦ Learn how data mining allows us to find patterns and regularities in databases
- ♦ Learn to apply the principles of data mining to the analysis of large complex datasets (Big Data), including those in very large databases or on web pages
- ♦ Explore, analyze and leverage data and convert it into useful and valuable information for clinical practice

### Module 3. Techniques for Extracting Genomic Data

- ♦ Understand how most scientific data appear in documents such as web pages and PDF files that are difficult to process for further analysis, however, using scraping techniques they can be used to
- ♦ Access to many data sources through the web for the implementation of precision medicine by allowing massive extraction of information

### Module 4. Application of Bioinformatics in Genomic Oncology

- ♦ Understanding the function of genes with little clinical information based on ontological proximity
- ♦ Discover genes involved in a disease based on a massive Pubmed search and graphical representation of the level of scientific evidence



*Improve the care of your patients by taking advantage of the specialization offered by the Genomic Oncology program"*

# 04 Course Management

The program includes in its teaching staff reference specialists in Machine Learning Techniques and Data Mining in Genomic Oncology who pour into this program the experience of their work. In addition, other specialists of recognized prestige participate in its design and elaboration, completing the program in an interdisciplinary manner.





“

*Learn from leading professionals the latest advances in Machine Learning and Data Mining Techniques in Genomic Oncology”*

## Management



### Dr. Oruezábal Moreno, Mauro Javier

- Head of the medical Oncology Service at La Paz University Hospital since 2017
- Research Visitors at University of Southampton
- Master's Degree in Bioinformatics and Biostatistics UOC-UB
- Master's Degree in bioinformatics analysis by the University Pablo de Olavide
- Doctor of Medicine from the Complutense University of Madrid. Outstanding Cum Laude Qualification
- Member of the Spanish Society of Medical Oncology and GECP Group (Spanish Group of Lung Cancer)
- Specialist (MIR) in Medical Oncology, University Hospital San Carlos of Madrid
- Degree in Medicine and Surgery, Navarra University



### Mr. Krallinger, Martin

- Head of the text mining unit at the Spanish National Cancer Research Center (CNIO)
- He has completed the selection process for the position of head of the text mining unit at the Barcelona Supercomputing Center (BSC)
- Expert in the field of biomedical and clinical text mining and linguistic technologies
- Expert in specific text mining applications for drug safety, molecular systems biology and oncology
- Participated in the implementation and evaluation of biomedical named entity recognition components, information extraction systems, semantic indexing of large datasets of heterogeneous document types
- Participated in the development of the first biomedical text annotation meta-server (biocreative metaserver - BCMS) and the BeCalm metaserver
- Organizer of BioCreative community evaluation challenges for the evaluation of natural language processing tools and has participated in the organization of biomedical text mining tasks in various international community challenges, including IberEval and CLEF

## Professors

### Mr. Alberich Martí, Ricardo

- ♦ Full university professor of Mathematics, Sciences and Computing (Director)
- ♦ Computer Science and Artificial Intelligence University of the Balearic Islands

### Ms. Álvarez Cubero, María Jesús

- ♦ Professor of the Department of Biochemistry III and Immunology, University of Granada

### Mr. Andrés León, Eduardo

- ♦ Head of the Bioinformatics Unit at the Institute of Parasitology and Biomedicine "Lopez-Neyra" - CSIC
- ♦ Degree in Biology and Molecular Biology from the Universidad Autónoma de Madrid

### Ms. Astudillo González, Aurora

- ♦ Anatomic Pathology Service
- ♦ Associate Professor at the University of Oviedo linked to the Central University Hospital of Asturias. Scientific Director of the Principality of Asturias Biobank

### Ms. Burón Fernández, María del Rosario

- ♦ Department of Internal Medicine, Infanta Cristina University Hospital

### Mr. Carmona Bayonas, Alberto

- ♦ Medical Oncology Service, Morales Meseguer General University Hospital

### Ms. Ciruelos, Eva M

- ♦ MD, Ph. D. Medical Oncology Service, 12 de Octubre University Hospital, Madrid
- ♦ HM CIOCC, Madrid

### Mr. Galiana, Enrique de Andrés

- ♦ Departments of Mathematics, University of Oviedo

### Mr. De la Haba Rodríguez, Juan

- ♦ Medical Oncology Department, University of Córdoba, Reina Sofía University Hospital, Córdoba, Spain

### Mr. Fernández Martínez, Juan Luis

- ♦ Director of the Inverse Problems, Optimization and Machine Learning Group, Department of Mathematics. University of Oviedo

### Ms. Figueroa, Angelica

- ♦ Institute of Biomedical Research A Coruña (INIBIC)
- ♦ Research Group Leader, Epithelial Plasticity and Metastasis

### Ms. García Casado, Zaida

- ♦ Laboratory of Molecular Biology Valencian Institute of Oncology Foundation

### Mr. García Foncillas, Jesús

- ♦ Jiménez Díaz, Foundation Medical Oncology Service

### Mr. Gomila Salas, Juan Gabriel

- ♦ University Professor Mathematical Sciences and Computer Science and Artificial Intelligence, University of the Balearic Islands

### Mr. González Gomáriz, José

- ♦ IdiSNA (Institute for Health Research of Navarra) Researcher in Training





**Mr. Hoyos Simón, Sergio**

- ♦ Medical Oncology Service at Rey Juan Carlos University Hospital

**Mr. Intxaurreondo, Ander**

- ♦ Life Sciences-Text Mining
- ♦ Barcelona Supercomputing Center

**Ms. Jiménez-Fonseca, Paula**

- ♦ Coordinator of the Digestive and Endocrine Tumors Section Medical Oncology. Asturias Central University Hospital

**Ms. Lage Alfranca, Yolanda**

- ♦ Jiménez Diaz, Foundation Medical Oncology Service

**Mr. López Guerrero, José Antonio**

- ♦ Medical Oncology Service, Valencian Institute of Oncology

**Mr. López López, Rafael**

- ♦ Head of the Medical Oncology Department
- ♦ Santiago de Compostela University Hospital Complex
- ♦ Translational Medical Oncology Group Health Research Institute

**Mr. Martínez González, Luis Javier**

- ♦ PhD Genomics Unit
- ♦ Pfizer center - University of Granada - Andalucía Government Center for Genomics and Oncology Research
- ♦ Pfizer - University of Granada - Andalucía Government Centre for Genomics and Oncological Research (GENYO)

**Ms. Martínez Iglesias, Olaia**

- ♦ Institute of Biomedical Research A Coruña (INIBIC)
- ♦ Research Group Leader, Epithelial Plasticity and Metástasis

**Mr. Paramio Gonzalez, Jesús María**

- ♦ CIEMAT Molecular Oncology Unit
- ♦ 12 de Octubre Research Institute of Madrid

**Mr. Pascual Martínez, Tomás**

- ♦ Barcelona Clinical Hospital
- ♦ Translational Genomics and Targeted Therapeutics in Solid Tumours Lab (IDIBAPS)

**Ms. Pérez Gutiérrez, Ana María**

- ♦ Student of the Master's Degree in the Clinical Bioinformatics Department of the Progress and Health Foundation -FPS- (Virgen del Rocío Hospital, Seville)
- ♦ PhD student in Biomedicine, UGR

**Ms. Ribalta, Teresa**

- ♦ MD, PhD. Chief, Anatomic Pathology Service, Hospital Sant Joan de Déu, Biobank
- ♦ Consultor, Anatomic Pathology Service, Hospital Clínic
- ♦ Professor of Pathology, Universitat de Barcelona

**Mr. Sánchez Rubio, Javier**

- ♦ Pharmacy Unit at Getafe University Hospital





**Mr. Olivas Varela, José Ángel**

- ◆ Deputy Director, Department of Information Technologies and Systems, Higher School of Computer Science

**Mr. Torres, Arnau Mir**

- ◆ Full university Professor of Mathematical Sciences and Computer Science and Artificial Intelligence, University of the Balearic Islands

**Mr. Soares, Felipe**

- ◆ Artificial Intelligence and Machine Learning Engineer at Apple
- ◆ Text Mining Research Engineer at the National Supercomputing Center in Barcelona

**Mr. Rueda Fernández, Daniel**

- ◆ Research Unit of the 12 de Octubre University Hospital, Madrid

**Mr. Segura Ruiz, Víctor**

- ◆ CIMA University of Navarra (Bioinformatics Platform) Unit Director

**Mr. Vázquez García, Miguel**

- ◆ Genome Informatics Group Leader
- ◆ Barcelona Supercomputing Center

**Mr. Velastegui Ordoñez, Alejandro**

- ◆ Medical Oncology Department

05

# Structure and Content

The structure of the contents has been designed by a team of professionals from the best educational centers, universities, and companies in the national territory, aware of the relevance of current specialization in order to intervene in the training and support of students, and committed to quality teaching through new educational technologies



Manufacturing  
Supply chain  
Product  
Cargo  
Customer  
Delivery  
Inventory  
Management  
Freight

Big  
Solution  
Marketing  
Analysis  
Ideas  
Success  
Mana

“

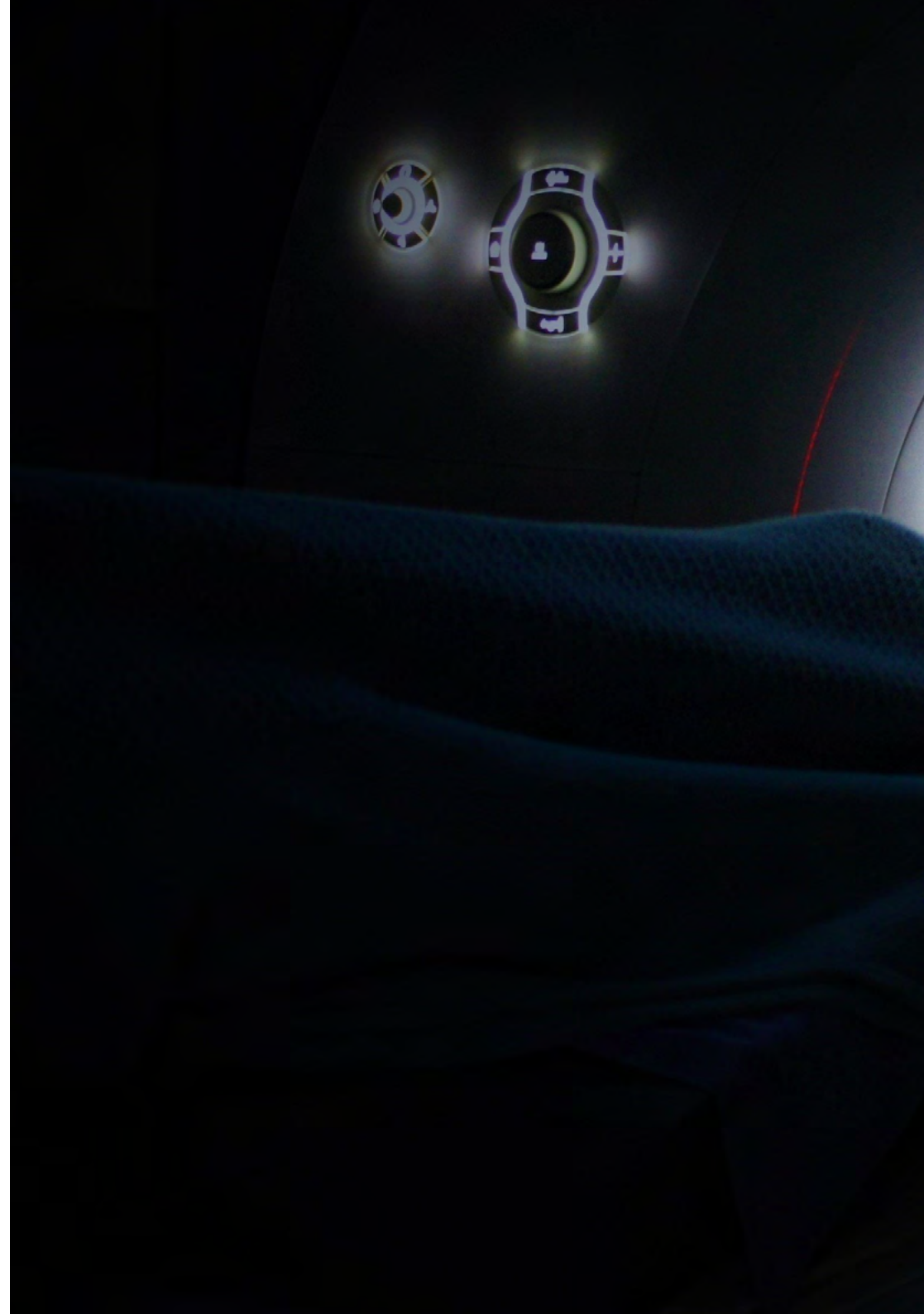
*This Postgraduate Diploma in Machine Learning and Data Mining Techniques in Genomic Oncology contains the most complete and up-to-date scientific program on the market”*

## Module 1. Machine learning for Analysing Big Data

- 1.1. Introduction to Machine Learning
- 1.2. Presentation of the Problem, Loading Data and Libraries
- 1.3. Data Cleaning (NAS, Categories, Dummy Variables)
- 1.4. Exploratory Data Analysis (Ggplot) + Crossed Validation
- 1.5. Prediction Algorithms: Multiple Linear Regression, Support Vector Machine, Regression Trees, Random Forest, etc.
- 1.6. Classification Algorithms: Logistic Regression, Support Vector Regression, Classification Trees, Random Forest
- 1.7. Adjustment of the Algorithm's Hyperparameters
- 1.8. Predicting Data with Different Models
- 1.9. ROC Curves and Confusion Matrices for Assessing Model Quality

## Module 2. Data Mining Applied to Genomics

- 2.1. Introduction
- 2.2. Initiation to Variables
- 2.3. Text Cleaning and Conditioning
- 2.4. Generating the Words Matrix
  - 2.4.1. Creating the TDM Words Matrix
  - 2.4.2. Visualizations on the TDM Word Matrix
- 2.5. Description of the Word Matrix
  - 2.5.1. Graphic Representation of the Frequencies
  - 2.5.2. Creating a Word Cloud
- 2.6. Creating a Data Frame for K-NN
- 2.7. Creating a Classification Model
- 2.8. Validating a Classification Model
- 2.9. Guided Practical Exercise on Data Mining in Cancer Genomics





### Module 3. Techniques for Extracting Genomic Data

- 3.1. Introduction to 'Scraping Data'
- 3.2. Importing Spreadsheet Data Files Stored Online
- 3.3. Scraping HTML Text
- 3.4. Scraping Data from an HTML Table
- 3.5. Using APIs for Data Scraping
- 3.6. Extracting Relevant Information
- 3.7. Using the Rvest Package of R
- 3.8. Obtaining Data Distributed Over Multiple Pages
- 3.9. Extracting Genomic Data from the "My Cancer Genome" Platform
- 3.10. Extracting Information on Genes from the "HGNC Hugo Gene Nomenclature Committee" Database
- 3.11. Extracting Pharmacological Data from the 'ONCOKG' (Precision Oncology Knowledge Base) Database

### Module 4. Application of Bioinformatics in Genomic Oncology

- 4.1. Clinical and Pharmacological Enrichment of Gene Variants
- 4.2. Mass Search in PubMed for Genomic Information
- 4.3. Mass Search in DGIdb for Genomic Information
- 4.4. Mass Search in Clinical Trials for Clinical Trials on Genomic Data
- 4.5. Gene Similarity Search for the Interpretation of a Gene Panel or Exome
- 4.6. Mass Search for Genes Connected to a Disease
- 4.7. Enrich-Gen: Platform for the Clinical and Pharmacological Enrichment of Genes
- 4.8. Procedure to Produce a Genomic Report in the Age of Precision Oncology

05

# Methodology

This academic program offers students a different way of learning. Our methodology uses a cyclical learning approach: **Relearning.**

This teaching system is used, for example, in the most prestigious medical schools in the world, and major publications such as the **New England Journal of Medicine** have considered it to be one of the most effective.





“

*Discover Relearning, a system that abandons conventional linear learning, to take you through cyclical teaching systems: a way of learning that has proven to be extremely effective, especially in subjects that require memorization"*

## At TECH we use the Case Method

What should a professional do in a given situation? Throughout the program, students will face multiple simulated clinical cases, based on real patients, in which they will have to do research, establish hypotheses, and ultimately resolve the situation. There is an abundance of scientific evidence on the effectiveness of the method. Specialists learn better, faster, and more sustainably over time.

*With TECH you will experience a way of learning that is shaking the foundations of traditional universities around the world.*



According to Dr. Gérvas, the clinical case is the annotated presentation of a patient, or group of patients, which becomes a "case", an example or model that illustrates some peculiar clinical component, either because of its teaching power or because of its uniqueness or rarity. It is essential that the case is based on current professional life, trying to recreate the real conditions in the physician's professional practice.

“

*Did you know that this method was developed in 1912, at Harvard, for law students? The case method consisted of presenting students with real-life, complex situations for them to make decisions and justify their decisions on how to solve them. In 1924, Harvard adopted it as a standard teaching method”*

The effectiveness of the method is justified by four fundamental achievements:

1. Students who follow this method not only achieve the assimilation of concepts, but also a development of their mental capacity, through exercises that evaluate real situations and the application of knowledge.
2. Learning is solidly translated into practical skills that allow the student to better integrate into the real world.
3. Ideas and concepts are understood more efficiently, given that the example situations are based on real-life.
4. Students like to feel that the effort they put into their studies is worthwhile. This then translates into a greater interest in learning and more time dedicated to working on the course.



## Relearning Methodology

At TECH we enhance the case method with the best 100% online teaching methodology available: Relearning.

This university is the first in the world to combine the study of clinical cases with a 100% online learning system based on repetition, combining a minimum of 8 different elements in each lesson, a real revolution with respect to the mere study and analysis of cases.

*Professionals will learn through real cases and by resolving complex situations in simulated learning environments. These simulations are developed using state-of-the-art software to facilitate immersive learning.*



At the forefront of world teaching, the Relearning method has managed to improve the overall satisfaction levels of professionals who complete their studies, with respect to the quality indicators of the best online university (Columbia University).

With this methodology, more than 250,000 physicians have been trained with unprecedented success in all clinical specialties regardless of surgical load. Our pedagogical methodology is developed in a highly competitive environment, with a university student body with a strong socioeconomic profile and an average age of 43.5 years old.

*Relearning will allow you to learn with less effort and better performance, involving you more in your specialization, developing a critical mindset, defending arguments, and contrasting opinions: a direct equation to success.*

In our program, learning is not a linear process, but rather a spiral (learn, unlearn, forget, and relearn). Therefore, we combine each of these elements concentrically.

The overall score obtained by TECH's learning system is 8.01, according to the highest international standards.



This program offers the best educational material, prepared with professionals in mind:



### Study Material

All teaching material is produced by the specialists who teach the course, specifically for the course, so that the teaching content is highly specific and precise.

These contents are then applied to the audiovisual format, to create the TECH online working method. All this, with the latest techniques that offer high quality pieces in each and every one of the materials that are made available to the student.



### Surgical Techniques and Procedures on Video

TECH introduces students to the latest techniques, the latest educational advances and to the forefront of current medical techniques. All of this in direct contact with students and explained in detail so as to aid their assimilation and understanding. And best of all, you can watch the videos as many times as you like.



### Interactive Summaries

The TECH team presents the contents attractively and dynamically in multimedia lessons that include audio, videos, images, diagrams, and concept maps in order to reinforce knowledge.

This exclusive educational system for presenting multimedia content was awarded by Microsoft as a "European Success Story".



### Additional Reading

Recent articles, consensus documents and international guidelines, among others. In TECH's virtual library, students will have access to everything they need to complete their course.





#### Expert-Led Case Studies and Case Analysis

Effective learning ought to be contextual. Therefore, TECH presents real cases in which the expert will guide students, focusing on and solving the different situations: a clear and direct way to achieve the highest degree of understanding.



#### Testing & Retesting

We periodically evaluate and re-evaluate students' knowledge throughout the program, through assessment and self-assessment activities and exercises, so that they can see how they are achieving their goals.



#### Classes

There is scientific evidence on the usefulness of learning by observing experts. The system known as Learning from an Expert strengthens knowledge and memory, and generates confidence in future difficult decisions.



#### Quick Action Guides

TECH offers the most relevant contents of the course in the form of worksheets or quick action guides. A synthetic, practical, and effective way to help students progress in their learning.



# 06 Certificate

The Postgraduate Diploma in Machine Learning and Data Mining Techniques in Genomic Oncology guarantees you, in addition to the most rigorous and updated training, access to a Postgraduate Diploma issued by TECH Technological University.





“

*Successfully complete this program  
and receive your university degree  
without travel or laborious paperwork”*

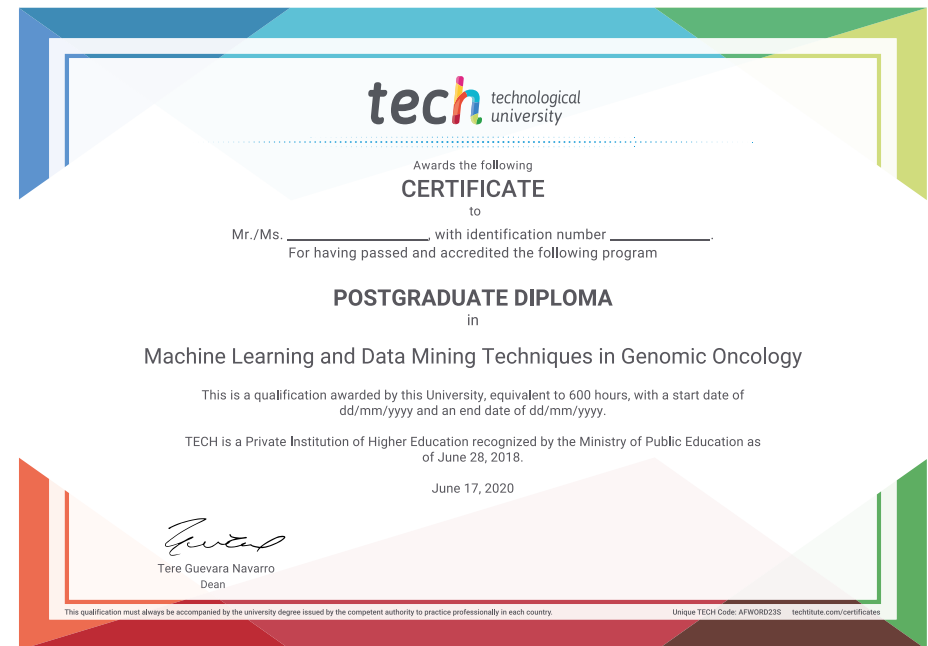
This **Postgraduate Diploma in Machine Learning and Data Mining Techniques in Genomic Oncology** contains the most complete and up-to-date scientific program on the market.

After the student has passed the evaluations, they will receive their corresponding **Postgraduate Diploma** issued by **TECH Technological University** via tracked delivery\*.

The certificate issued by **TECH Technological University** will reflect the qualification obtained in the Postgraduate Diploma, and meets the requirements commonly demanded by labor exchanges, competitive examinations, and professional career evaluation committees.

Title: **Postgraduate Diploma in Machine Learning and Data Mining Techniques in Genomic Oncology**

Official N° of hours: **600 h.**



\*Apostille Convention. In the event that the student wishes to have their paper certificate issued with an apostille, TECH EDUCATION will make the necessary arrangements to obtain it, at an additional cost.

health future confidence people education information tutors guarantee accreditation teaching institutions technology learning community commitment personalized service innovation knowledge present development languages virtual classroom



## Postgraduate Diploma

Machine Learning  
and Data Mining  
Techniques in  
Genomic Oncology

Course Modality: Online

Duration: 6 months

Certificate: TECH Technological University

Official N° of hours: 600 h.

# Postgraduate Diploma

Machine Learning and Data Mining  
Techniques in Genomic Oncology