

# Universitätskurs

## Verarbeitung und Transformation in Data Mining



## Universitätskurs Verarbeitung und Transformation in Data Mining

- » Modalität: **online**
- » Dauer: **6 Wochen**
- » Qualifizierung: **TECH Technische Universität**
- » Aufwand: **16 Std./Woche**
- » Zeitplan: **in Ihrem eigenen Tempo**
- » Prüfungen: **online**

Internetzugang: [www.techtute.com/de/informatik/universitatskurs/verarbeitung-transformation-data-mining](http://www.techtute.com/de/informatik/universitatskurs/verarbeitung-transformation-data-mining)

# Index

01

Präsentation

---

Seite 4

02

Ziele

---

Seite 8

03

Kursleitung

---

Seite 12

04

Struktur und Inhalt

---

Seite 16

05

Methodik

---

Seite 20

06

Qualifizierung

---

Seite 28

# 01

# Präsentation

*Data Mining* wird in einer Analyse vor dem Einsatz von Techniken des maschinellen Lernens in der *Data Science* vereinfacht. Letzteres hilft Computeringenieuren, den größtmöglichen Nutzen aus den Daten zu ziehen. Im Rahmen des Programms werden die Studenten die verschiedenen statistischen Inferenzverfahren erlernen, um die Untersuchung von Variablen, deren Verteilung und Werte zu analysieren und zu bestimmen. Dies ist ein Karriereaufstieg für Studenten, die sich in diesem Bereich spezialisieren möchten.



“

*Entwickeln Sie Ihre Fähigkeiten  
als Informatiker in Data Science  
und Data Mining"*

In diesem Universitätskurs werden die theoretischen Grundlagen analysiert, die Computingingenieuren helfen, fortgeschrittene Kenntnisse über die verschiedenen bestehenden Datenaufbereitungstechniken zur Datenbereinigung, -normalisierung und -transformation zu entwickeln. Darüber hinaus werden die Instrumente vorgestellt werden, die erforderlich sind, um verschiedene Methoden auf Fehler zu untersuchen, die in der Arbeitsumgebung Probleme verursachen können.

Das gesamte Programm besteht aus einer Reihe von Fallstudien, die den Lernprozess von Studenten unterstützen, die ihre Karriere vorantreiben und sich selbst herausfordern wollen, um Spitzenleistungen zu erzielen.

All dies wird dank eines 100%igen Online-Programms, das sich an die täglichen Bedürfnisse der Studenten anpasst, greifbar. Es ist nur ein Gerät mit Internetanschluss erforderlich, um mit der Arbeit für ein komplettes berufliches Profil mit internationaler Ausstrahlung zu beginnen.

Dieser **Universitätskurs in Verarbeitung und Transformation in Data Mining** enthält das vollständigste und aktuellste Programm auf dem Markt. Die herausragendsten Merkmale der Spezialisierung sind:

- ♦ Die Entwicklung von Fallstudien, die von technischen Experten im Bereich der Datenanalyse vorgestellt werden
- ♦ Der anschauliche, schematische und äußerst praxisnahe Inhalt soll wissenschaftliche und praktische Informationen zu den für die berufliche Praxis wesentlichen Disziplinen vermitteln
- ♦ Praktische Übungen, bei denen der Selbstbewertungsprozess zur Verbesserung des Lernens genutzt werden kann
- ♦ Sein besonderer Schwerpunkt liegt auf innovativen Methoden
- ♦ Theoretische Vorträge, Fragen an den Experten, Diskussionsforen zu kontroversen Themen und individuelle Reflexionsarbeit
- ♦ Die Verfügbarkeit des Zugriffs auf die Inhalte von jedem festen oder tragbaren Gerät mit Internetanschluss



*Bewerten Sie die  
verschiedenen vorgestellten  
Methoden und ermitteln Sie  
die Vor- und Nachteile"*



*Entwickeln Sie die für die Identifizierung, Aufbereitung und Umwandlung von Daten erforderlichen Fähigkeiten"*

Zu den Dozenten des Programms gehören Fachleute aus der Branche, die ihre Erfahrungen aus ihrer Arbeit in diese Fortbildung einbringen, sowie anerkannte Spezialisten aus führenden Unternehmen und renommierten Universitäten.

Die multimedialen Inhalte, die mit der neuesten Bildungstechnologie entwickelt wurden, werden der Fachkraft ein situiertes und kontextbezogenes Lernen ermöglichen, d. h. eine simulierte Umgebung, die eine immersive Fortbildung bietet, die auf das Training in realen Situationen ausgerichtet ist.

Das Konzept dieses Studiengangs konzentriert sich auf problemorientiertes Lernen, bei dem die Fachkraft versuchen wird, die verschiedenen Situationen aus der beruflichen Praxis zu lösen, die während des akademischen Programms auftreten. Dies wird mit Hilfe eines innovativen Systems interaktiver Videos geschehen, die von renommierten Experten mit umfassender Erfahrung in Verarbeitung und Transformation in *Data Mining* erstellt wurden.

*Konkretisieren Sie wirksame und effiziente Verfahren zur Datenverarbeitung je nach Art der Problemstellung.*

*Krempeln Sie Ihre Karriere um und beginnen Sie mit der Entwicklung von Verbesserungsstrategien in einem Unternehmen.*



# 02 Ziele

Das in diesem Programm vermittelte Wissen wird Computeringenieuren dabei helfen, die Analyse vor der Anwendung von Techniken des maschinellen Lernens anzugehen, um den größtmöglichen Nutzen aus den Daten zu ziehen. So werden sie in der Lage sein, Fachwissen auf der Grundlage statistischer Analysen zu generieren. Zu diesem Zweck hat TECH die folgenden allgemeinen und spezifischen Ziele festgelegt.





“

*Demonstrieren Sie Ihre  
Fähigkeit, Datenvisualisierungen  
für beschreibende Analysen  
zu interpretieren"*



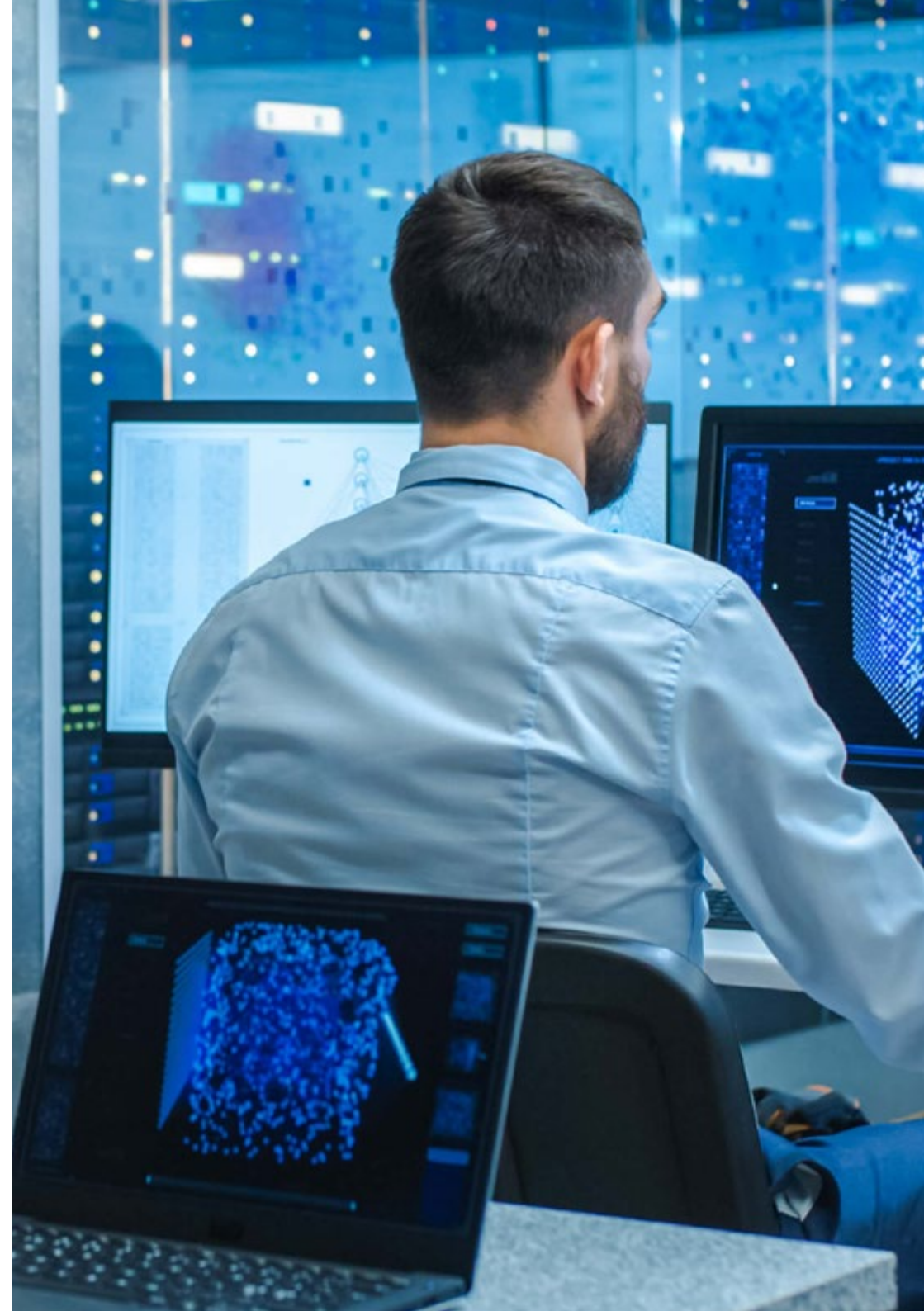
## Allgemeine Ziele

---

- ◆ Untersuchen der Vorteile der Anwendung von Datenanalysetechniken in jeder Abteilung des Unternehmens
- ◆ Entwickeln der Grundlage für das Verständnis der Bedürfnisse und Anwendungen der einzelnen Abteilungen
- ◆ Generieren von Fachwissen, um das richtige Werkzeug auszuwählen
- ◆ Vorschlagen von Techniken und Zielen, um je nach Abteilung so produktiv wie möglich zu sein



*Untersuchen Sie die Probleme in Ihrem Umfeld und finden Sie praktikable Lösungen, indem Sie Daten als Grundlage für Ihre Vorschläge verwenden"*





## Spezifische Ziele

---

- ◆ Generieren von Fachwissen über die vorherige Statistik für die Datenanalyse und -auswertung
- ◆ Entwickeln der notwendigen Fähigkeiten zur Identifizierung, Vorbereitung und Umwandlung von Daten
- ◆ Bewerten der verschiedenen vorgestellten Methoden und Identifizieren der Vor- und Nachteile
- ◆ Untersuchen von Problemen in hochdimensionalen Datenumgebungen
- ◆ Entwickeln der Implementierung der Algorithmen für die Datenvorverarbeitung
- ◆ Demonstrieren der Fähigkeit, Datenvisualisierungen für die deskriptive Analyse zu interpretieren
- ◆ Entwickeln fortgeschrittener Kenntnisse über die verschiedenen vorhandenen Datenaufbereitungstechniken zur Datenbereinigung, Normalisierung und Datentransformation



# 03 Kursleitung

Der Universitätskurs in Verarbeitung und Transformation in *Data Mining* bringt eine ausgewählte Gruppe von Fachleuten mit mehrjähriger Erfahrung in der Datenanalyse im Unternehmenssektor zusammen. Auf diese Weise wird sichergestellt, dass das vermittelte Wissen von Fachleuten stammt, die in der Lage sind, alle Fragen der Studenten zu beantworten und ihnen reale Fälle vorzulegen, um den Inhalt des Programms zu veranschaulichen.





“

*Mit ihrer mehrjährigen Erfahrung in diesem Bereich wird diese Gruppe von Experten Ihnen jederzeit helfen, die Konzepte des Programms auf einfache und didaktische Weise zu verstehen"*

## Leitung



### Dr. Peralta Martín-Palomino, Arturo

- ♦ CEO und CTO bei Prometheus Global Solutions
- ♦ CTO bei Korporate Technologies
- ♦ CTO bei AI Shephers GmbH
- ♦ Promotion in technischer Informatik an der Universität von Castilla La Mancha
- ♦ Promotion in Wirtschaftswissenschaften, Unternehmen und Finanzen an der Universität Camilo José Cela
- ♦ Außerordentlicher Promotionspreis
- ♦ Promotion in Psychologie an der Universität von Castilla La Mancha
- ♦ Masterstudiengang in Fortgeschrittene Informationstechnologien an der Universität von Castilla La Mancha
- ♦ Masterstudiengang MBA+E (Master in Business Administration and Organisational Engineering) an der Universität von Castilla La Mancha
- ♦ Außerordentlicher Professor für Bachelor- und Masterstudiengänge in Computertechnik der an der Universität von Castilla La Mancha
- ♦ Professor für den Masterstudiengang in Big Data und Datenwissenschaft an der Internationalen Universität von Valencia
- ♦ Professor für den Masterstudiengang in Industrie 4.0 und den Masterstudiengang in Industriedesign und Produktentwicklung
- ♦ Mitglied der SMILe-Forschungsgruppe der Universität von Castilla La Mancha



## Professoren

### Hr. Montoro Montarroso, Andrés

- ◆ Forscher in der SMILe-Gruppe an der Universität von Castilla La Mancha
- ◆ Datenwissenschaftler bei Prometheus Global Solutions
- ◆ Hochschulabschluss in Computertechnik an der Universität von Castilla La Mancha
- ◆ Masterstudiengang in Datenwissenschaft und Computertechnik an der Universität von Granada
- ◆ Gastprofessor für das Fach Wissensbasierte Systeme an der Hochschule für Informatik in Ciudad Real, für die Vorlesung: "Fortgeschrittene Techniken der künstlichen Intelligenz: Suche und Analyse potenzieller Radikaler in den sozialen Medien"
- ◆ Gastprofessor für das Fach Data Mining an der Hochschule für Informatik in Ciudad Real, für die Vorlesung: "Anwendungen der Verarbeitung natürlicher Sprache: Fuzzy Logic zur Analyse von Nachrichten in sozialen Netzwerken"
- ◆ Referent auf dem Seminar über Korruptionsprävention in öffentlichen Verwaltungen und künstliche Intelligenz an der Fakultät für Rechts- und Sozialwissenschaften von Toledo, für die Konferenz mit dem Titel "Techniken der künstlichen Intelligenz"
- ◆ Referent auf dem ersten internationalen Seminar über Verwaltungsrecht und künstliche Intelligenz (DAIA), organisiert von dem Europäischen Studienzentrums Luis Ortega Álvarez und dem Forschungsinstitut TransJus, für die Konferenz mit dem Titel "Analyse von Gefühlen zur Verhinderung von Hassreden in sozialen Medien"

# 04

# Struktur und Inhalt

Die Module dieses Programms bieten eine theoretische und praktische Perspektive, um modernste Datenbereinigungstechniken, Transformation, Dimensionalitätsreduktion, *Feature*- und Instanzauswahl zu untersuchen. So wird das Ziel des Programms, professionelle, umfassende und hoch angesehene Ingenieure zu qualifizieren, erreicht.



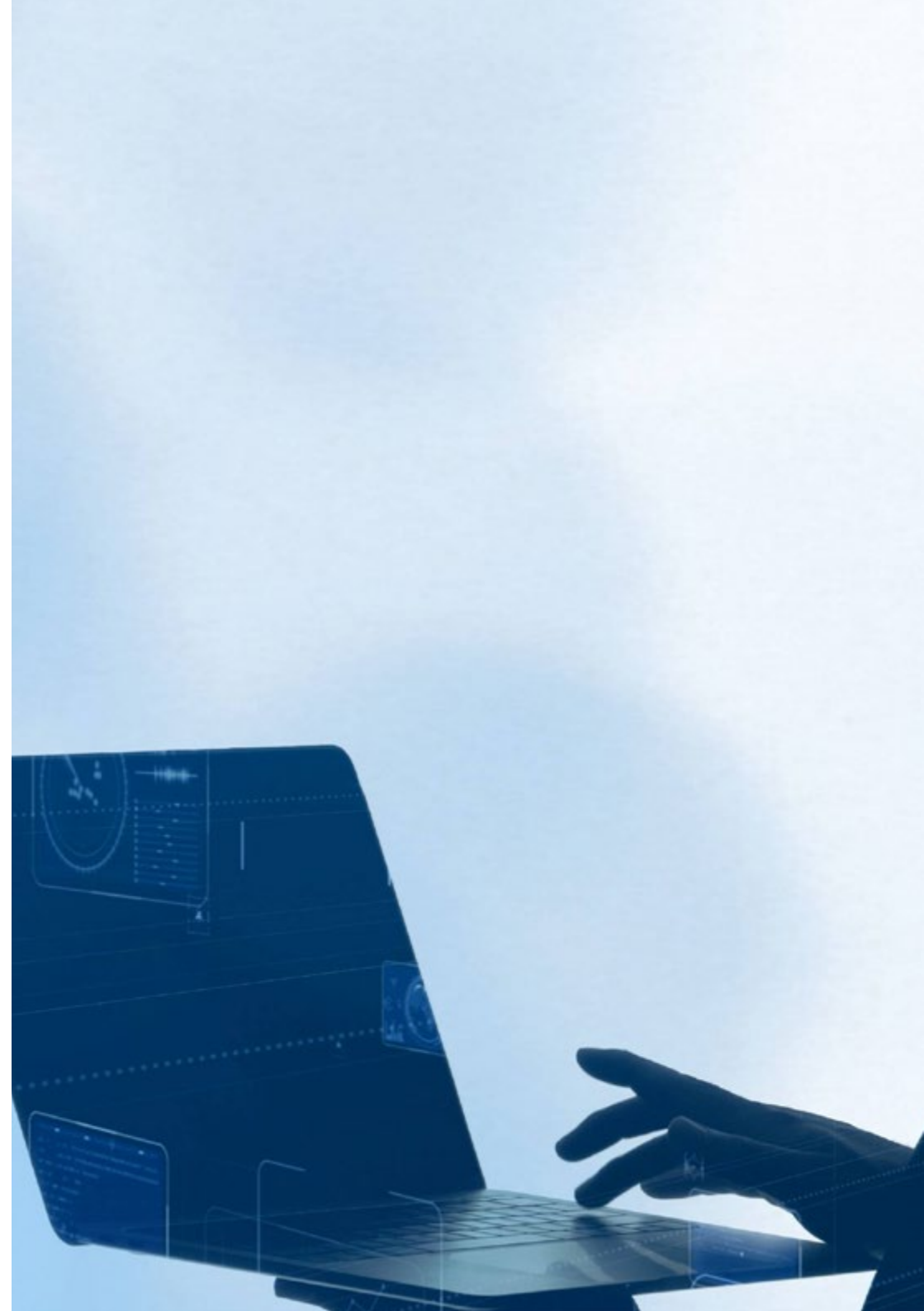




“Entwickeln Sie fortgeschrittene Kenntnisse über die verschiedenen vorhandenen Datenaufbereitungstechniken zur Datenbereinigung, -normalisierung und -umwandlung”

## Modul 1. Data Mining. Auswahl, Vorverarbeitung und Transformation

- 1.1. Statistische Inferenz
  - 1.1.1. Deskriptive Statistik vs. Statistische Inferenz
  - 1.1.2. Parametrische Verfahren
  - 1.1.3. Nicht-parametrische Verfahren
- 1.2. Explorative Analyse
  - 1.2.1. Deskriptive Analyse
  - 1.2.2. Visualisierung
  - 1.2.3. Vorbereitung der Daten
- 1.3. Vorbereitung der Daten
  - 1.3.1. Datenintegration und -bereinigung
  - 1.3.2. Normalisierung der Daten
  - 1.3.3. Attribute umwandeln
- 1.4. Verlorene Werte
  - 1.4.1. Umgang mit verlorenen Werten
  - 1.4.2. *Maximum-Likelihood*-Imputationsmethoden
  - 1.4.3. Imputation verlorener Werte durch maschinelles Lernen
- 1.5. Datenrauschen
  - 1.5.1. Lärmklassen und Attribute
  - 1.5.2. Rauschfilterung
  - 1.5.3. Rauscheffekt
- 1.6. Der Fluch der Dimensionalität
  - 1.6.1. *Oversampling*
  - 1.6.2. *Undersampling*
  - 1.6.3. Multidimensionale Datenreduktion
- 1.7. Kontinuierliche zu diskreten Attributen
  - 1.7.1. Kontinuierliche versus diskrete Daten
  - 1.7.2. Prozess der Diskretisierung



- 1.8. Daten
  - 1.8.1. Datenauswahl
  - 1.8.2. Perspektiven und Auswahlkriterien
  - 1.8.3. Methoden der Auswahl
- 1.9. Auswahl der Instanzen
  - 1.9.1. Methoden für die Instanzauswahl
  - 1.9.2. Auswahl der Prototypen
  - 1.9.3. Erweiterte Methoden für die Instanzauswahl
- 1.10. Vorverarbeitung von Daten in *Big Data*-Umgebungen
  - 1.10.1. *Big Data*
  - 1.10.2. "Klassische" versus massive Vorbearbeitung
  - 1.10.3. *Smart Data*



*Der Erwerb dieses  
Programms wird den  
Studenten ein besseres  
Verständnis der Methoden  
der Datenauswahl vermitteln"*

# 05 Methodik

Dieses Fortbildungsprogramm bietet eine andere Art des Lernens. Unsere Methodik wird durch eine zyklische Lernmethode entwickelt: **das Relearning**.

Dieses Lehrsystem wird z. B. an den renommiertesten medizinischen Fakultäten der Welt angewandt und wird von wichtigen Publikationen wie dem **New England Journal of Medicine** als eines der effektivsten angesehen.



“

*Entdecken Sie Relearning, ein System, das das herkömmliche lineare Lernen aufgibt und Sie durch zyklische Lehrsysteme führt: eine Art des Lernens, die sich als äußerst effektiv erwiesen hat, insbesondere in Fächern, die Auswendiglernen erfordern"*

## Fallstudie zur Kontextualisierung aller Inhalte

Unser Programm bietet eine revolutionäre Methode zur Entwicklung von Fähigkeiten und Kenntnissen. Unser Ziel ist es, Kompetenzen in einem sich wandelnden, wettbewerbsorientierten und sehr anspruchsvollen Umfeld zu stärken.

“

*Mit TECH werden Sie eine Art des Lernens erleben, die die Grundlagen der traditionellen Universitäten in der ganzen Welt verschiebt”*



*Sie werden Zugang zu einem Lernsystem haben, das auf Wiederholung basiert, mit natürlichem und progressivem Unterricht während des gesamten Lehrplans.*



*Die Studenten lernen durch gemeinschaftliche Aktivitäten und reale Fälle die Lösung komplexer Situationen in realen Geschäftsumgebungen.*

## Eine innovative und andersartige Lernmethode

Dieses TECH-Programm ist ein von Grund auf neu entwickeltes, intensives Lehrprogramm, das die anspruchsvollsten Herausforderungen und Entscheidungen in diesem Bereich sowohl auf nationaler als auch auf internationaler Ebene vorsieht. Dank dieser Methodik wird das persönliche und berufliche Wachstum gefördert und ein entscheidender Schritt in Richtung Erfolg gemacht. Die Fallmethode, die Technik, die diesem Inhalt zugrunde liegt, gewährleistet, dass die aktuellste wirtschaftliche, soziale und berufliche Realität berücksichtigt wird.

**“** *Unser Programm bereitet Sie darauf vor, sich neuen Herausforderungen in einem unsicheren Umfeld zu stellen und in Ihrer Karriere erfolgreich zu sein“*

Die Fallmethode ist das am weitesten verbreitete Lernsystem an den besten Informatikschulen der Welt, seit es sie gibt. Die Fallmethode wurde 1912 entwickelt, damit die Jurastudenten das Recht nicht nur anhand theoretischer Inhalte erlernen, sondern ihnen reale, komplexe Situationen vorlegen, damit sie fundierte Entscheidungen treffen und Werturteile darüber fällen können, wie diese zu lösen sind. Sie wurde 1924 als Standardlehrmethode in Harvard eingeführt.

Was sollte eine Fachkraft in einer bestimmten Situation tun? Mit dieser Frage konfrontieren wir Sie in der Fallmethode, einer handlungsorientierten Lernmethode. Während des gesamten Kurses werden die Studierenden mit mehreren realen Fällen konfrontiert. Sie müssen Ihr gesamtes Wissen integrieren, recherchieren, argumentieren und Ihre Ideen und Entscheidungen verteidigen.

## Relearning Methodik

TECH kombiniert die Methodik der Fallstudien effektiv mit einem 100%igen Online-Lernsystem, das auf Wiederholung basiert und in jeder Lektion verschiedene didaktische Elemente kombiniert.

Wir ergänzen die Fallstudie mit der besten 100%igen Online-Lehrmethode: Relearning.

*Im Jahr 2019 erzielten wir die besten  
Lernergebnisse aller spanischsprachigen  
Online-Universitäten der Welt.*

Bei TECH lernen Sie mit einer hochmodernen Methodik, die darauf ausgerichtet ist, die Führungskräfte der Zukunft auszubilden. Diese Methode, die an der Spitze der weltweiten Pädagogik steht, wird Relearning genannt.

Unsere Universität ist die einzige in der spanischsprachigen Welt, die für die Anwendung dieser erfolgreichen Methode zugelassen ist. Im Jahr 2019 ist es uns gelungen, die Gesamtzufriedenheit unserer Studenten (Qualität der Lehre, Qualität der Materialien, Kursstruktur, Ziele...) in Bezug auf die Indikatoren der besten Online-Universität in Spanisch zu verbessern.





In unserem Programm ist das Lernen kein linearer Prozess, sondern erfolgt in einer Spirale (lernen, verlernen, vergessen und neu lernen). Daher wird jedes dieser Elemente konzentrisch kombiniert. Mit dieser Methode wurden mehr als 650.000 Hochschulabsolventen mit beispiellosem Erfolg in so unterschiedlichen Bereichen wie Biochemie, Genetik, Chirurgie, internationales Recht, Managementfähigkeiten, Sportwissenschaft, Philosophie, Recht, Ingenieurwesen, Journalismus, Geschichte, Finanzmärkte und -Instrumente ausgebildet. Dies alles in einem sehr anspruchsvollen Umfeld mit einer Studentenschaft mit hohem sozioökonomischem Profil und einem Durchschnittsalter von 43,5 Jahren.

*Das Relearning ermöglicht es Ihnen, mit weniger Aufwand und mehr Leistung zu lernen, sich mehr auf Ihr Fachgebiet einzulassen, einen kritischen Geist zu entwickeln, Argumente zu verteidigen und Meinungen zu kontrastieren: eine direkte Gleichung zum Erfolg.*

Nach den neuesten wissenschaftlichen Erkenntnissen der Neurowissenschaften wissen wir nicht nur, wie wir Informationen, Ideen, Bilder und Erinnerungen organisieren, sondern auch, dass der Ort und der Kontext, in dem wir etwas gelernt haben, von grundlegender Bedeutung dafür sind, dass wir uns daran erinnern und es im Hippocampus speichern können, um es in unserem Langzeitgedächtnis zu behalten.

Auf diese Weise sind die verschiedenen Elemente unseres Programms im Rahmen des so genannten neurokognitiven kontextabhängigen E-Learnings mit dem Kontext verbunden, in dem der Teilnehmer seine berufliche Praxis entwickelt.



Dieses Programm bietet die besten Lehrmaterialien, die sorgfältig für Fachleute aufbereitet sind:



#### Studienmaterial

Alle didaktischen Inhalte werden von den Fachleuten, die den Kurs unterrichten werden, speziell für den Kurs erstellt, so dass die didaktische Entwicklung wirklich spezifisch und konkret ist.

Diese Inhalte werden dann auf das audiovisuelle Format angewendet, um die TECH-Online-Arbeitsmethode zu schaffen. Und das alles mit den neuesten Techniken, die dem Studenten qualitativ hochwertige Stücke aus jedem einzelnen Material zur Verfügung stellen.



#### Meisterklassen

Die Nützlichkeit der Expertenbeobachtung ist wissenschaftlich belegt.

Das sogenannte Learning from an Expert baut Wissen und Gedächtnis auf und schafft Vertrauen für zukünftige schwierige Entscheidungen.



#### Fertigkeiten und Kompetenzen Praktiken

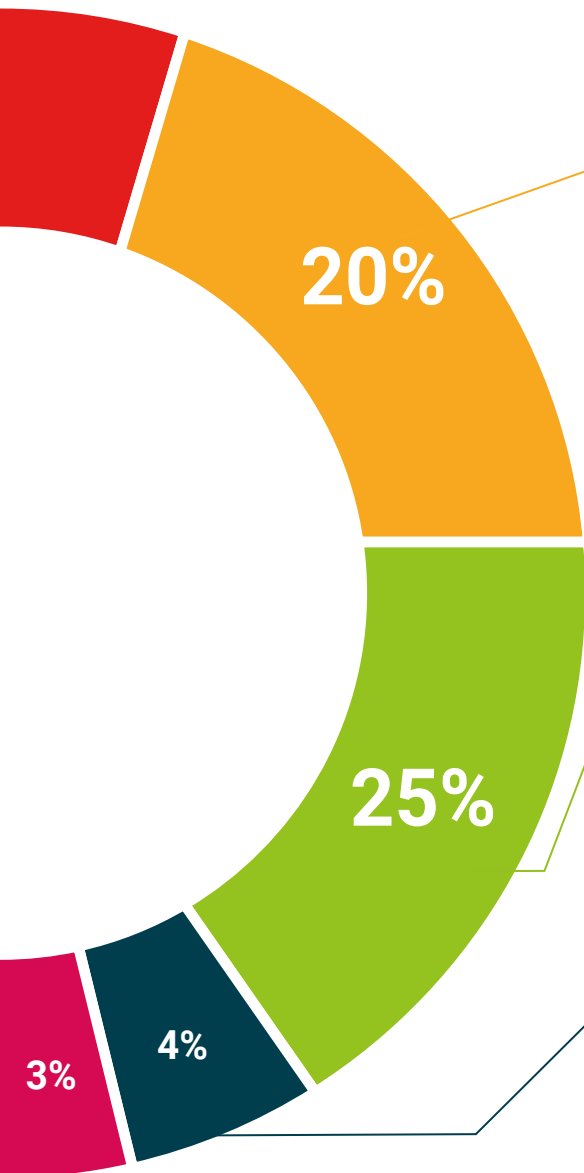
Sie werden Aktivitäten durchführen, um spezifische Kompetenzen und Fertigkeiten in jedem Fachbereich zu entwickeln. Praktiken und Dynamiken zum Erwerb und zur Entwicklung der Fähigkeiten und Fertigkeiten, die ein Spezialist im Rahmen der Globalisierung, in der wir leben, entwickeln muss.



#### Weitere Lektüren

Aktuelle Artikel, Konsensdokumente und internationale Leitfäden, u.a. In der virtuellen Bibliothek von TECH haben die Studenten Zugang zu allem, was sie für ihre Ausbildung benötigen.





#### Fallstudien

Sie werden eine Auswahl der besten Fallstudien vervollständigen, die speziell für diese Qualifizierung ausgewählt wurden. Die Fälle werden von den besten Spezialisten der internationalen Szene präsentiert, analysiert und betreut.



#### Interaktive Zusammenfassungen

Das TECH-Team präsentiert die Inhalte auf attraktive und dynamische Weise in multimedialen Pillen, die Audios, Videos, Bilder, Diagramme und konzeptionelle Karten enthalten, um das Wissen zu vertiefen.

Dieses einzigartige Bildungssystem für die Präsentation multimedialer Inhalte wurde von Microsoft als "europäische Erfolgsgeschichte" ausgezeichnet.



#### Prüfung und Nachprüfung

Die Kenntnisse der Studenten werden während des gesamten Programms regelmäßig durch Bewertungs- und Selbsteinschätzungsaktivitäten und -übungen beurteilt und neu bewertet, so dass die Studenten überprüfen können, wie sie ihre Ziele erreichen.



06

# Qualifizierung

Der Universitätskurs in Verarbeitung und Transformation in Data Mining garantiert neben der strengsten und aktuellsten Ausbildung auch den Zugang zu einem von der TECH Technologischen Universität ausgestellten Diplom.



“

*Schließen Sie dieses Programm  
erfolgreich ab und erhalten Sie  
Ihren Universitätsabschluss ohne  
lästige Reisen oder Formalitäten”*

Dieser **Universitätskurs in Verarbeitung und Transformation in Data Mining** enthält das vollständigste und aktuellste Programm auf dem Markt.

Sobald der Student die Prüfungen bestanden hat, erhält er/sie per Post\* mit Empfangsbestätigung das entsprechende Diplom, ausgestellt von der **TECH Technologischen Universität**.

Das von **TECH Technologische Universität** ausgestellte Diplom drückt die erworbene Qualifikation aus und entspricht den Anforderungen, die in der Regel von Stellenbörsen, Auswahlprüfungen und Berufsbildungsausschüssen verlangt werden.

Titel: **Universitätskurs in Verarbeitung und Transformation in Data Mining**

Anzahl der offiziellen Arbeitsstunden: **150 Std.**



zukunft

gesundheit vertrauen menschen  
erziehung information tutoren  
garantie akkreditierung unterricht  
institutionen technologie lernen

**tech** technologische  
universität

Universitätskurs  
Verarbeitung und  
Transformation  
in Data Mining

- » Modalität: online
- » Dauer: 6 Wochen
- » Qualifizierung: TECH Technologische Universität
- » Aufwand: 16 Std./Woche
- » Zeitplan: in Ihrem eigenen Tempo
- » Prüfungen: online

# Universitätskurs

## Verarbeitung und Transformation in Data Mining

```
ngSwitchWatchAction(val  
elements.length; i < ii; ++i)  
remove();  
length = 0;  
selectedScopes.length; i < ii; ++i) {  
selected = selectedElements[i];  
selectedScopes[i].$destroy();  
previousElements[i] = selected;  
animate.leave(selected, function() {  
previousElements.splice(i, 1);  
});  
}
```

```
selectedElements.length = 0;  
selectedScopes.length = 0;
```

```
if ((selectedTranscludes = ngSwitchController.selectedTranscludes))
```