

# 大学课程

## 数据挖掘处理和转换



## 大学课程 数据挖掘处理和转换

- » 模式: 在线
- » 时间: 6周
- » 学位: TECH科技大学
- » 时间: 16小时/周
- » 时间表: 按你方便的
- » 考试: 在线

网络访问: [www.techitute.com/cn/information-technology/postgraduate-certificate/data-mining-processing-transformation](http://www.techitute.com/cn/information-technology/postgraduate-certificate/data-mining-processing-transformation)

# 目录

01

介绍

---

4

02

目标

---

8

03

课程管理

---

12

04

结构和内容

---

18

05

方法

---

22

06

学位

---

30

# 01 介绍

在数据科学中使用机器学习技术之前，数据挖掘在分析中被简化。后者有助于计算机工程师从数据中提取最大价值。通过该课程，学生将学习不同的统计推断程序，以分析和确定变量、变量分布和变量值的研究。对于希望在这一领域有所专长的学生来说，这是一种职业发展。



“

在数据科学和数据挖掘领域  
培养计算机工程师的技能”

在本大学课程中,将分析有助于计算机工程师发展有关数据清理、规范化和转换等不同现有数据准备技术的高级知识的理论基础。它还将介绍评估不同方法的必要工具,以找出可能在工作环境中造成问题的错误。

整个课程由一系列案例研究组成,这些案例研究将为那些寻求进一步发展职业生涯和挑战自我、追求卓越的学生提供学习支持。

这个100%的在线课程,可适应了学生的日常需要,只需要有一个有互联网连接的设备,就可以学习,以获得一个完整、具有国际影响力的专业形象。

这个**数据挖掘处理和转换大学课程**包含市场上最完整和最新的课程。主要特点是:

- ◆ 由工程专家介绍的案例研究的发展侧重于数据分析
- ◆ 该书的内容图文并茂、示意性强、实用性强为那些视专业实践至关重要的学科提供了科学和实用的信息
- ◆ 可以进行自我评估过程的实践,以推进学习
- ◆ 其特别强调创新方法
- ◆ 理论课、向专家提问、关于有争议问题的讨论区和个人反思性论文
- ◆ 可以从任何有互联网连接的固定或便携式设备上获取内容



评估提出的不同方法并确定优缺点”



“培养识别、准备和转换数据的必要技能”

根据提出的问题类型,将有效和高效的数据处理程序具体化。

扭转你的职业生涯,开始为公司制定改进战略。

该课程的教学人员包括来自该行业的专业人士,他们将自己的工作经验带到了这一培训中,还有来自领先公司和著名大学的公认专家。

多媒体内容是用最新的教育技术开发的,将允许专业人员进行情景式学习,即一个模拟的环境,提供一个身临其境的培训,为真实情况进行培训。

该课程的设计重点是基于问题的学习,通过这种方式,专业人员必须尝试解决整个课程中出现的不同专业实践情况。这将借助由在数据挖掘处理和转换方面拥有丰富经验的知名专家制作的创新型互动视频系统来实现。



# 02 目标

本课程提供的知识将帮助计算机工程师在应用机器学习技术之前进行分析,以便从数据中获取最大价值。这样,他们就能在统计分析的基础上生成专业知识。为此,TECH研究所制定了以下总体目标和。







“

展示你解释数据可视化以  
进行描述性分析的能力”



## 总体目标

- ◆ 分析在公司各个部门应用数据分析技术的好处
- ◆ 开发了解每个部门的需求和应用的基础
- ◆ 学习专业知识以选择正确的工具
- ◆ 根据部门提出尽可能高效的技术和目标



研究你所处环境中的问题, 并以数据为基础找到可行的解决方案"





## 具体目标

---

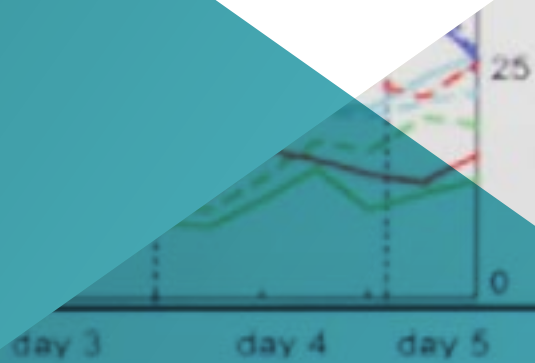
- ◆ 为任何数据分析和评估发展统计数据的专业知识
- ◆ 培养识别、准备和转换数据的必要技能
- ◆ 评估提出的不同方法并确定优缺点
- ◆ 检查高维数据环境的问题
- ◆ 开发用于数据预处理的算法
- ◆ 展示解释数据可视化以进行描述性分析的能力
- ◆ 开发有关用于数据清理、规范和转换的不同现有数据准备技术的高级知识





# 03 课程管理

数据挖掘处理和转换大学课程汇集了一批在商业领域拥有多年数据分析经验的专业人士。这确保了传授知识的专业人士能够回答学生可能提出的任何问题，并为他们提供真实案例，以更好地体现课程内容。





“

这批专家拥有多年的行业经验,他们将随时帮助你以轻松和说教的方式理解课程的概念”



## 国际客座董事

Tom Flowerdew 博士是数据科学领域享有国际声誉的杰出人物。他曾在伦敦担任万事达卡的数据科学副总裁。在这个角色中,他负责该领域团队的准备、运营和策略,旨在支持一系列创新支付产品、打击反洗钱 (AML) 并分析加密货币的使用案例。

此外,他还在万事达卡的网络情报解决方案部门担任数据科学总监,领导数据整合工作以支持基于加密货币的革命性产品。事实上,他处理复杂数据和开发先进解决方案的能力在多个网络安全和金融领域的项目中取得了关键成功。

同样地,他在 Featurespace 公司担任多个重要角色,包括标准化产品交付主管,在剑桥领导一个团队并推动转型项目,将交付时间和精力减少了75%以上。此外,作为美国地区交付总监,他管理了公司在北美的所有交付职能,显著提升了运营效率并加强了与客户的关系。

另外, Tom Flowerdew 博士在他的职业生涯中展示了建立和领导高绩效团队的能力,尤其是在亚特兰大担任数据科学家时,他招聘并管理了一组该领域的专家团队,以及在剑桥。他的创新和解决问题的方式在他工作过的组织中留下了深刻的印记,使他成为数据科学领域的有影响力的领导者。



## Flowerdew, Tom 博士

---

- 万事达卡数据科学副总裁, 伦敦, 英国
- 万事达卡网络情报解决方案数据科学总监, 伦敦
- Featurespace 标准化产品交付主管, 剑桥
- Featurespace 美国地区交付总监, 剑桥
- Featurespace 数据科学家, 亚特兰大, 乔治亚州, 美国
- Featurespace 数据科学家, 剑桥
- 兰卡斯特大学统计与运筹学研究员
- 兰卡斯特大学运筹学博士
- BAE Systems 系统工程学士
- 约克大学数学学士

“

感谢 TECH, 你将能够与世界上最优秀的专业人士一起学习”

## 管理人员



### Peralta Martín-Palomino, Arturo 博士

- Prometheus Global Solutions 的CEO和CTO
- Korporate Technologies的首席技术官
- AI Shepherds GmbH 首席技术官
- 卡斯蒂利亚拉曼恰大学计算机工程博士
- Camilo José Cela 大学经济学、商业和金融学博士。非凡博士奖
- 卡斯蒂利亚拉曼恰大学心理学博士
- 卡斯蒂利亚拉曼恰大学高级信息技术硕士
- 卡斯蒂利亚拉曼恰大学MBA+E (工商管理 and 组织工程硕士)
- 卡斯蒂利亚拉曼恰大学计算机工程学士和硕士学位的副教授
- 瓦伦西亚国际大学大数据和数据科学硕士的教授
- 工业 4.0 硕士和工业设计与产品开发硕士的教授
- 卡斯蒂利亚拉曼恰大学 SMILe 研究小组成员



## 教师

### Montoro Montarroso, Andrés先生

- ◆ 卡斯蒂利亚-拉曼恰大学 SMILe 小组研究员
- ◆ Prometheus Global Solutions 的数据科学家
- ◆ 卡斯蒂利亚拉曼恰大学计算机工程学士
- ◆ 格拉纳达大学数据科学与计算机工程硕士
- ◆ Escuela Superior de Informática de Ciudad Real 的知识系统主题客座教授, 发表演讲: “先进的人工智能技术: 搜索和分析社交媒体中的潜在激进分子”
- ◆ Escuela Superior de Informática de Ciudad Real 数据挖掘主题的客座教授, 在会议上发表演讲: “自然语言处理的应用: 社交网络中消息分析的模糊逻辑”
- ◆ 在公共行政和人工智能预防腐败研讨会上发言。托莱多法律和社会科学学院。发表题目为“人工智能技术”。第一届行政法与人工智能国际研讨会 (DAIA) 的演讲者。由 Luis Ortega Álvarez 欧洲研究中心和 TransJus 研究所组织。发表题目: “预防社交网络中仇恨信息的情感分析”

# 04

## 结构和内容

该课程的各个模块从理论和实践的角度出发,研究最先进的数据清理技术、转换、降维、特征和实例选择。这样,培养专业、全面和高声望工程师的计划目标就得以实现。





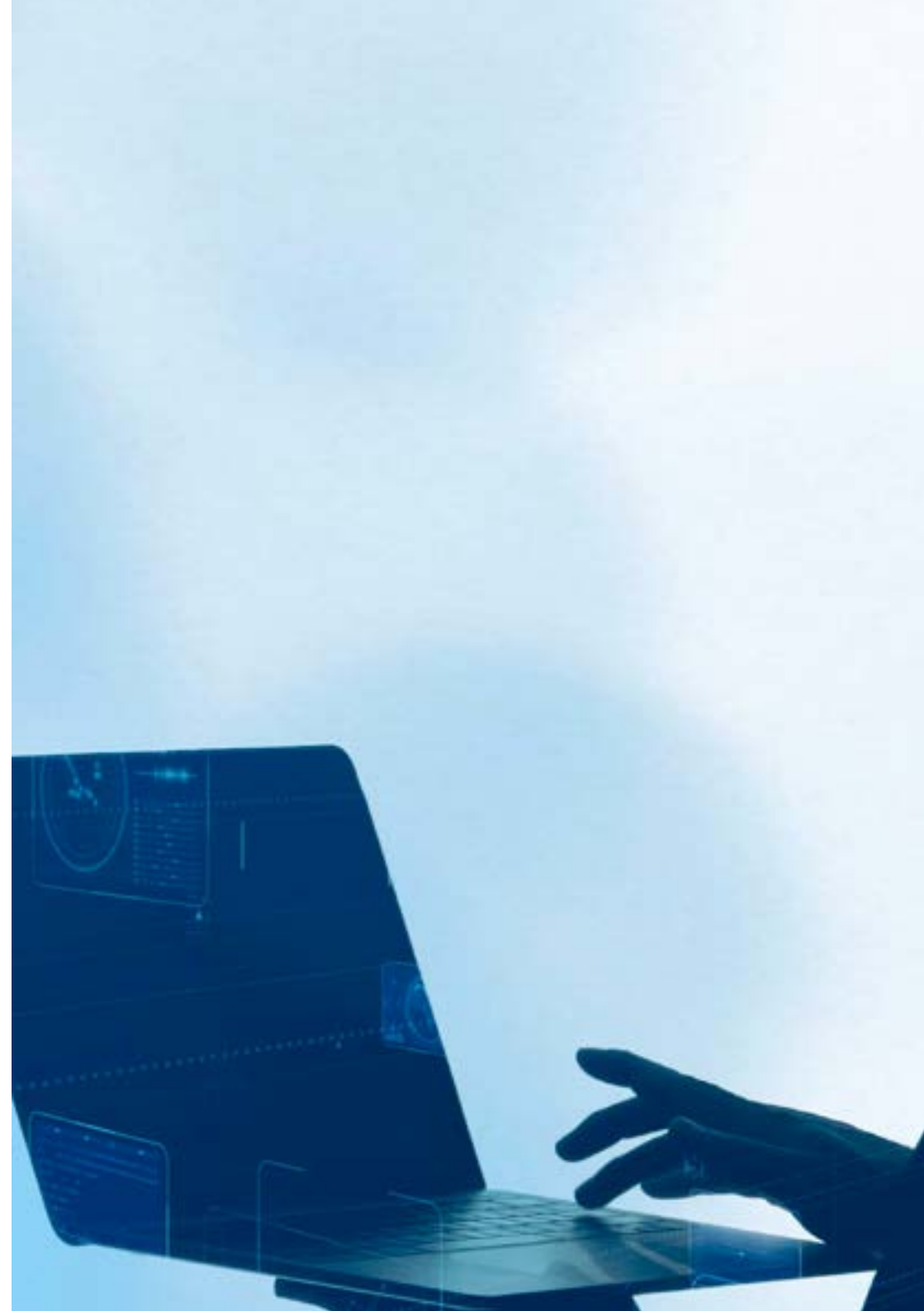


“

开发有关用于数据清理、规范和转换的不同现有数据准备技术的高级知识”

## 模块1.数据挖掘选择、预处理和转换

- 1.1. 统计推断
  - 1.1.1. 描述性统计和统计推断
  - 1.1.2. 参数化程序
  - 1.1.3. 非参数过程
- 1.2. 探索性分析
  - 1.2.1. 描述性分析
  - 1.2.2. 视觉化
  - 1.2.3. 数据准备
- 1.3. 数据准备
  - 1.3.1. 数据整合和清理
  - 1.3.2. 数据标准化
  - 1.3.3. 转换属性
- 1.4. 缺失值
  - 1.4.1. 缺失值的处理
  - 1.4.2. 最大似然插补方法
  - 1.4.3. 使用机器学习估算缺失值
- 1.5. 数据中的噪音
  - 1.5.1. 噪声类别和属性
  - 1.5.2. 噪声过滤
  - 1.5.3. 噪音的影响
- 1.6. 维度的祸害
  - 1.6.1. 过度采样
  - 1.6.2. 采样不足
  - 1.6.3. 多维数据缩减
- 1.7. 从连续属性到离散属性
  - 1.7.1. 连续数据与离散数据
  - 1.7.2. 离散化过程



- 1.8. 数据
  - 1.8.1. 数据选取
  - 1.8.2. 观点和选择标准
  - 1.8.3. 挑选方法
- 1.9. 选择阶段
  - 1.9.1. 选择阶段的方法
  - 1.9.2. 原型的选择
  - 1.9.3. 选择阶段的高级方法
- 1.10. 大数据环境中的数据预处理
  - 1.10.1. 大数据
  - 1.10.2. “经典”与批量预处理
  - 1.10.3. 智能数据



完成该课程将使学生更好地了解数据选择方法"



# 05 方法

这个培训计划提供了一种不同的学习方式。我们的方法是通过循环的学习模式发展起来的：**再学习**。

这个教学系统被世界上一些最著名的医学院所采用，并被**新英格兰医学杂志**等权威出版物认为是最有效的教学系统之一。







“

发现再学习, 这个系统放弃了传统的线性学习, 带你体验循环教学系统: 这种学习方式已经证明了其巨大的有效性, 尤其是在需要记忆的科目中”



## 案例研究, 了解所有内容的背景

我们的方案提供了一种革命性的技能和知识发展方法。我们的目标是在一个不断变化, 竞争激烈和高要求的环境中加强能力建设。

“

和TECH, 你可以体验到一种正在动摇  
世界各地传统大学基础的学习方式”



你将进入一个以重复为基础的学习系统, 在  
整个教学大纲中采用自然和渐进式教学。



学生将通过合作活动和真实案例，学习如何解决真实商业环境中的复杂情况。

### 一种创新并不同的学习方法

该技术课程是一个密集的教学计划，从零开始，提出了该领域在国内和国际上最苛刻的挑战和决定。由于这种方法，个人和职业成长得到了促进，向成功迈出了决定性的一步。案例法是构成这一内容的技术基础，确保遵循当前经济、社会和职业现实。

“我们的课程使你准备好在不确定的环境中面对新的挑战，并取得事业上的成功”

在世界顶级计算机科学学校存在的时间里，案例法一直是最广泛使用的学习系统。1912年开发的案例法是为了让法律学生不仅在理论内容的基础上学习法律，案例法向他们展示真实的复杂情况，让他们就如何解决这些问题作出明智的决定和价值判断。1924年，它被确立为哈佛大学的一种标准教学方法。

在特定情况下，专业人士应该怎么做？这就是我们在案例法中面对的问题，这是一种以行动为导向的学习方法。在整个课程中，学生将面对多个真实的案例。他们必须整合所有的知识，研究、论证和捍卫他们的想法和决定。

## 再学习方法

TECH有效地将案例研究方法基于循环的100%在线学习系统相结合,在每节课中结合了个不同的教学元素。

我们用最好的100%在线教学方法加强案例研究:再学习。

在2019年,我们取得了世界上所有西班牙语在线大学中最好的学习成绩。

在TECH,你将用一种旨在培训未来管理人员的尖端方法进行学习。这种处于世界教育学前沿的方法被称为再学习。

我校是唯一获准使用这一成功方法的西班牙语大学。2019年,我们成功地提高了学生的整体满意度(教学质量,材料质量,课程结构,目标.....),与西班牙语最佳在线大学的指标相匹配。





在我们的方案中,学习不是一个线性的过程,而是以螺旋式的方式发生(学习,解除学习,忘记和重新学习)。因此,我们将这些元素中的每一个都结合起来。这种方法已经培养了超过65万名大学毕业生,在生物化学,遗传学,外科,国际法,管理技能,体育科学,哲学,法律,工程,新闻,历史,金融市场和工具等不同领域取得了前所未有的成功。所有这些都是在一个高要求的环境中进行的,大学学生的社会经济状况很好,平均年龄为43.5岁。

再学习将使你的学习事半功倍,表现更出色,使你更多地参与到训练中,培养批判精神,捍卫论点和对比意见:直接等同于成功。

从神经科学领域的最新科学证据来看,我们不仅知道如何组织信息,想法,图像记忆,而且知道我们学到东西的地方和背景,这是我们记住并将其储存在海马体的根本原因,并能将其保留在长期记忆中。

通过这种方式,在所谓的神经认知背景依赖的电子学习中,我们课程的不同元素与学员发展其专业实践的背景相联系。





该方案提供了最好的教育材料,为专业人士做了充分准备:



### 学习材料

所有的教学内容都是由教授该课程的专家专门为该课程创作的,因此,教学的发展是具体的。

然后,这些内容被应用于视听格式,创造了TECH在线工作方法。所有这些,都是用最新的技术,提供最高质量的材料,供学生使用。



### 大师课程

有科学证据表明第三方专家观察的有用性。

向专家学习可以加强知识和记忆,并为未来的困难决策建立信心。



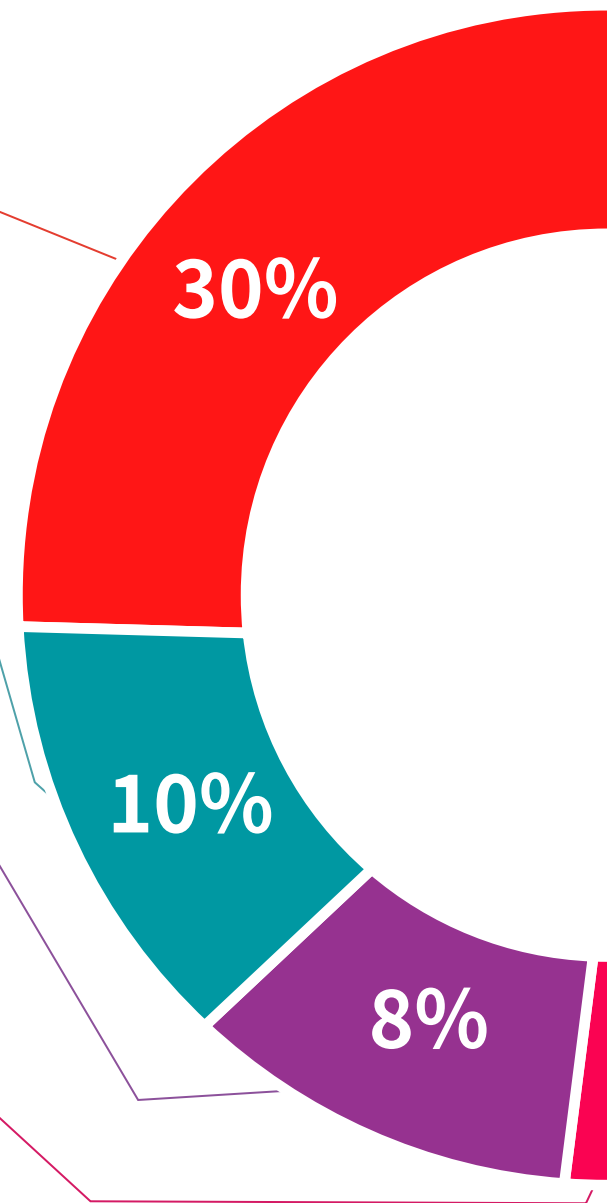
### 技能和能力的实践

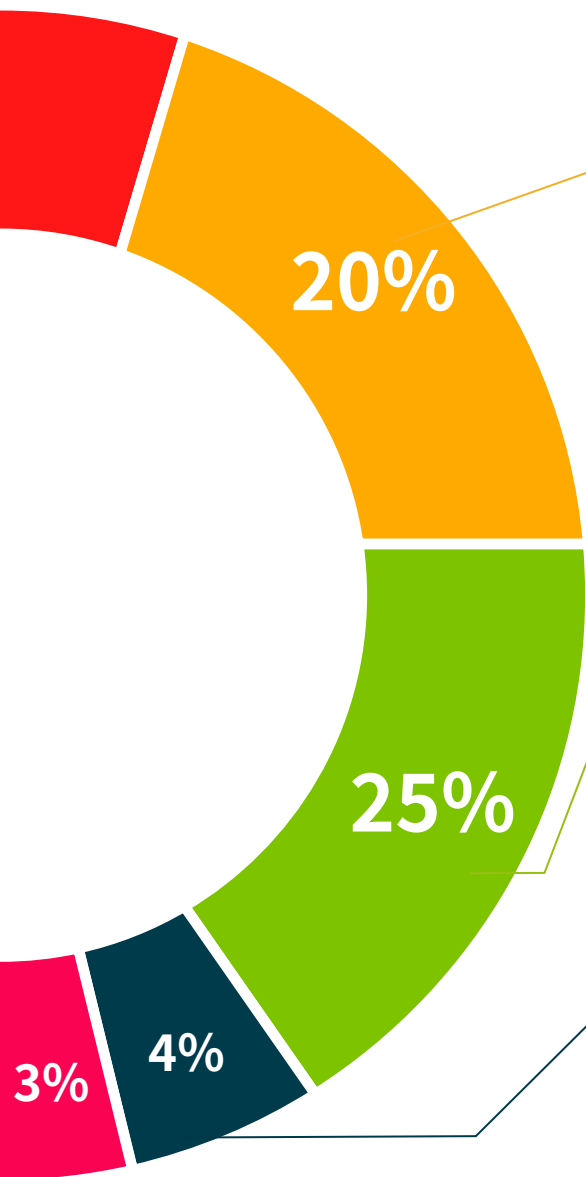
你将开展活动以发展每个学科领域的具体能力和技能。在我们所处的全球化框架内,我们提供实践和氛围帮你取得成为专家所需的技能和能力。



### 延伸阅读

最近的文章,共识文件和国际准则等。在TECH的虚拟图书馆里,学生可以获得他们完成培训所需的一切。





### 案例研究

他们将完成专门为这个学位选择的最佳案例研究。由国际上最好的专家介绍,分析和辅导案例。



### 互动式总结

TECH团队以有吸引力和动态的方式将内容呈现在多媒体中,其中包括音频,视频,图像,图表和概念图,以强化知识。  
这个用于展示多媒体内容的独特教育系统被微软授予“欧洲成功案例”称号。



### 测试和循环测试

在整个课程中,通过评估和自我评估活动和练习,定期评估和重新评估学习者的知识:通过这种方式,学习者可以看到他/她是如何实现其目标的。



# 06 学位

数据挖掘处理和转换大学课程除了保证最严格和最新的培训外,还可以获得由TECH科技大学颁发的大学课程学位证书。



“

顺利完成该课程并获得大学课程, 无需旅行或文书工作的麻烦”



这个**数据挖掘处理和转换大学课程**包含了市场上最完整和最新的课程。

评估通过后, 学生将通过邮寄收到**TECH科技大学**颁发的相应的**大学课程学位**。

**TECH科技大学**颁发的证书将表达在大学课程获得的资格, 并将满足工作交流, 竞争性考试和专业职业评估委员会的普遍要求。

学位:**数据挖掘处理和转换大学课程**

官方学时:**150小时**



健康 信心 未来 人 导师  
教育 信息 教学  
保证 资格认证 学习  
机构 社区 科技 承诺  
个性化的关注 现在 创新  
知识 网页 培 质量  
网上教室 发展 语言 机构

**tech** 科学技术大学

大学课程  
数据挖掘处理和转换

- » 模式:在线
- » 时间:6周
- » 学位:TECH科技大学
- » 时间:16小时/周
- » 时间表:按你方便的
- » 考试:在线

# 大学课程

## 数据挖掘处理和转换

```
ngSwitchWatchAction(valu  
elements.length; i < ii; ++i)  
move());  
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
```