

Business Intelligence

קורס בינה עסקית

תיאור הקורס

מסלול התמחות זה מכשיר את המועמד לתפקיד מפתח BI. השיעורים מורכבים מתרגולים מעמיקים על מנת לאפשר לסטודנט להתמודד בהצלחה עם האתגרים בעולם הבינה העסקית. תכנית הלימודים במסלול זה הינה ייחודית ומעודכנת לפי צרכי השוק והטכנולוגיה. ההכשרה כוללת שימוש בחבילת הנתונים של Microsoft - SSIS, SSAS - SQL Server.

קורס זה מתחלק לשלושה חלקים:

חלק ראשון - שפת SQL ו TSQL:

SQL היא שפה המאפשרת גישה ומניפולציה למידע השמור במסדי נתונים. בעזרת שאילתות בשפת SQL ניתן להציג מידע מבסיס הנתונים, לעדכן אותן, למחוק, להכניס מידע חדש ועוד. בנוסף ניתן ליצור בסיסי נתונים וטבלאות חדשות, ואף לשנות את המבנה שלהם או את האינדקס שלהם.

Transact-SQL או TSQL היא שפת פיתוח לבסיסי נתונים רלציוניים של מיקרוסופט ו-Sybase והיא מרחיבה במקרים רבים את SQL וכוללת יכולות שאינן כלולות בו. כל היכולות וההרחבות האלו הופכות את TSQL ל-Turing Complete.

חלק שני - בינה עסקית (Business Intelligence):

הבינה העסקית היא תחום בטכנולוגיית מידע העוסק בבניית מערכות העוזרות לארגון להפיק מידע עסקי משמעותי תוך שימוש נרחב בכלל הנתונים הנאספים על ידו מכלל מערכות הארגון השונות.

מערכות בינה עסקית מספקות מידע היסטורי לצד מידע עדכני ואל מול תחזיות הנוגעות לפעילות העסקית.

בעזרת נתונים אלה יכול מנהל הארגון לאתר דפוסים משמעותיים בניהול הארגון, לאתר סוגיות הדורשות טיפול, לאתר ליקויים, לבנות צפי לעתיד ולבצע ניתוח היסטורי של הנעשה בארגון.

ניתן לומר כי הבינה העסקית הינה ה"עיניים והאוזניים" של כל חברה גדולה. הנתונים מכילים את כל התשובות לכל השאלות, יש רק צורך לבנות תשתית מתאימה, לשאול את השאלות הנכונות ולדעת כיצד להפיק את התשובות.

חלק שלישי - למידת מכונה (Machine Learning):

למידת מכונה הינו ללא ספק תחום העתיד. תחום זה עוסק בבניית אלגוריתמים אשר אינם קובעים באופן דטרמיניסטי כיצד יש לבצע מטלה מסוימת כי אם מקנים למערכת יכולת לשפר את ביצוע המשימה ככל שהיא נחשפת לעוד ועוד נתונים. השימוש הגובר והולך בנתוני עתק (Big Data) מאפשר "אימון" של המכונה בכמות עצומה של קבוצות נתונים דבר המאפשר "ללמד" מכונות כיצד לבצע משימות אשר עד עתה נחשבו לאנושיות

חשוב לדעת

היקף השעות: 470 שעות אקדמאיות:

חלק א' – 60 ש"א.

חלק ב' – 200 ש"א + 25 שעות עבודה על פרוייקט.

חלק ג' – 160 ש"א + 25 שעות עבודה על פרוייקט.

קהל יעד ודרישות קבלה: מסלול הכשרה זה מיועד לבעלי רקע טכני במערכות מידע, מפתחים, בודקי תוכנה ואנשי תשתיות. נדרשת יכולת שילוב והבנת המידע עסקי תוך יישומו בטכנולוגיה, ניסיון בעבודה עם שאילתות SQL ואנגלית ברמה גבוהה.

מתודולוגיות למידה: קורס זה כולל תכנים מקצועיים ומקוריים של חברת **Microsoft** כמו כן הקורס מלווה בסרטונים ייעודיים של המכללה בתחום נתוני העתק (BIG DATA). בסיום הקורס, הסטודנטים יבצעו פרויקט מעשי המסכם ומשלב את כל הידע הנרכש בקורס בכל מסלול ההכשרה. מטרת הפרויקט הינה להתמחות בכלל הכלים, לצורך צבירת ניסיון מעשי ולרכוש ניסיון רב בפתרון תקלות העולות במהלך הפיתוח.

סגל המרצים: למכללת iNT טכנולוגיות סגל מרצים ומומחי הדרכה, מהמובילים בתחום, בעלי ניסיון מעשי רב ביישום והדרכת נושאי הלימוד בתעשיית ההי-טק הישראלית והעולמית.

זכאות לתעודת גמר מטעם מכללת iNT: תעודת גמר מטעם מכללת iNT תוענק לבוגרים העומדים בכל דרישות התוכנית לרבות נוכחות ב 85% מהשיעורים לפחות, הגשת המטלות ופרויקט הגמר בהתאם לדרישות הקורס ועמידה בתקנון הלימודים. בסיום הקורס יוכלו התלמידים לגשת לבחינות ההסמכה הבאות (תלמידים העוברים בהצלחה בחינות אלו יקבלו תעודת הסמכה רשמית ממיקרוסופט):

20461c - Querying Microsoft SQL Server

20463c - Implementing a Data Warehouse with Microsoft SQL Server

20466c - Implementing Data Models and Reports with Microsoft SQL Server

תכנית לימודים

חלק א'

לימוד שפת SQL ו TSQL (60 ש"א)

לימוד שפת SQL ו TSQL, רכישת מושגים מעולם ה DBA ומסד הנתונים, תרגול שפת השאילתות ברמה גבוהה. היחידה כוללת בנוסף הקניית הבנה בסיסית וידע במסדי נתונים, שימוש בשפת SQL ליצירת יחסי גומלין תוך לימוד Data Integrity. אנו נצלול את תוך עולם ה SQL ונבין כיצד ליצור קשרי גומלין בין טבלאיים וכיצד לבצע שאילתות מורכבות.

פלטפורמת BI – ליבת הקורס (200 ש"א)

(35 ש"א) Data Model Design

Dimensional Model - ביחידה זו נלמד כיצד לתכנן ולבנות Data Warehouse תוך שימוש ב SQL Server וכן בתוכנת מידול מקצועית – Erwin Data Modeler.

המודול יעסוק ביסודות ה Dimensional Model.

אנו נלמד אודות טבלאות FACT ו Dimensions , נבין את המשמעות של Measures ובמה הם נבדלים מ Dimensions, נלמד אודות מודל סכימת כוכב (Star Schema) כמקרה פרטי של מודל פתית שלג (Snowflake Schema) וכן נלמד כיצד לבצע מידול. בנוסף, נכיר מושגים חשובים מתחום ה - Dimensional Model כגון - Slowly Changing Dimensions of type 0, 1 and 2.

הלימוד ילווה בדוגמאות מעשיות המדמות מידול של Business Processes בתעשיות שונות. בסיום המודול יגישו הסטודנטים פרויקט מעשי המדמה מקרה הלקוח מהעולם העסקי.

SSIS – SQL Server Integration Service (60 ש"א)

כלי ETL המאפשר שליפת מידע ממקורות שונים וטעינתם למחסן נתונים – Warehouse Data – שלב גזירה, טרנספורמציה וטעינת הנתונים.

SAS Tabular Model (50 ש"א)

מודל הנתונים הטבולרי מציג גישה שונה מזו של המודל הרב מימדי. הוא שומר על הקשרים הבין טבלאיים ומפשט את מודל הקוביה. בהיותו מיועד עבור בסיסי נתונים שוכני זכרון הוא מאפשר הרצת שאילתות במהירות גבוהה מאוד ושליפה מהירה של נתונים. עבור מסדי נתונים שאינם גדולים מאוד יכול מודל זה להוות מודל אידיאלי לבניית DWH. ביחידה זו נלמד על מודל הנתונים הטבולרי וכן על שפת DAX המאפשרת ביצוע של ניתוח אנליטי במודלים אלו.

המודול כולל לימוד כלל פלטפורמת ה - BI של Microsoft תוך סקירה ולימוד של כליים נוספים.

(40 ש"א) Data Visualization Fundamentals

ביחידה זו נלמד את עקרונות ה Data Visualization, Similarity, Proximity, Gestalt Principles, Continuity, Closure & Connectedness, נלמד כיצד להמחיש את הנתונים וליצור גרפים, דשבורדים מורכבים ו Stories.

כמו כן נלמד לעבוד עם כלי מרכזי בתחום: Power BI וכן נלמד לחברו לאקסל, לאינטרנט, למסדי נתונים ובעצם לכל מקור נתונים נתמך שנחפוץ..

(15 ש"א) Big Data

היחידה האחרונה בקורס תתמקד בעולם Big Data ו Data science נכיר את המושגים המרכזיים והטכנולוגיות בהן נעשה שימוש בתעשייה כיום, בדגש על "גן החיות" של Apache, חלק זה של הקורס ישלב למידה מתוקשבת באמצעות סרטונים במערכת הלמידה של המכללה.

חלק ג'

(160 ש"א) Machine Learning

מבוא ל - Machine Learning (80 ש"א)

בחלק זה, נלמד את העקרונות באמצעותם אפשר ללמד את המחשב לפתור בעיות באמצעות בניית פונקציות מטרה והרצתה על סטים גדולים של נתונים. בנוסף, נלמד את השיטות והאלגוריתמים של למידה חישובית ושל זיהוי תבניות. המטרה המרכזית של למידת המכונה הינה בניית מערכת לומדת המחקר בהתנהגותה את המח האנושי.

מימוש מודלים של שפת מכונה (80 ש"א)

חלק זה של המסלול מתמקד בלימוד שפת מכונה ומקנה יכולות עשירות ומגוונות המצטרפות כנדבך נוסף מעל החלק הראשון המתמקד בהכשרת מפתח BI.

לימוד שפת SQL | TSQ (60 שעות)

1. Introduction
 - What is a Relational Data Base Management System?
 - Understanding the Power of Relations
 - Basic Concepts – Tables, fields
 - Introducing Microsoft SQL Server Management Studio
2. SQL Fundamentals
 - Select Statements
 - Where Clause & Filters
 - Logical Conditions and Predicates
 - Null Values and Null Functions
 - Data Types
 - DISTINCT
 - Order By
 - Group By
 - Aggregation functions
 - AVG
 - MAX
 - MIN
 - SUM
 - COUNT
 - Using Aliases
3. Mastering Relations
 - Basic Joins
 - Primary Key
 - Foreign Key
 - Types of Relations - Many to Many, One to Many, One to One
 - Converting Many to Many into a One to Many and Many to One relation
 - Left Join, Right Join, Inner Join
 - Nested Queries
 - Self Join – Joining different rows at the same table
4. SQL Programming
 - SQL Enjoin pint of view – Understanding the order of SQL commands as a key for more accurate and efficient SQL programming
 - System Functions
 - Effective Design and implementation of User Defined Functions
 - Effective Design and implementation of Stored Procedures
 - Using Parameters Effectively
 - Working with Triggers
 - Understanding Semi Structure data- XMLs, Json files (Note: this topic will be elaborated in the Big Data module)
 - Storing and Reading Json & XML files in tables
5. Constraints & Data Integrity
 - Understanding and Imposing Constrains
 - Unique
 - Not Null
 - PK & FK as constraints
 - Enforcing Data Integrity
 - Understanding Normalization
6. Schema Commands
 - Creating a DB
 - Creating tables
 - Modifying and enhancing the table schema
7. Data Base Design
 - Indexes
 - Types of Indexes
 - Index Impact over performance
 - Column Indexes
 - Partitioning
 - Views – Why and when to use Views
8. Data Manipulation
 - Inserting data into tables
 - Deleting data from tables
 - Updating table data
 - Duplicating table data

פלטפורמת BI – ליבת הקורס

Data Model Design

- An Introduction to Data Warehouse
 - What is a Data warehouse and why is it essential for BI projects?
 - Review and Understand Data hubs and alternative BI solutions
 - Data warehouse basic concepts
 - Considerations for an effective design and implementation of a Data warehouse
 - De normalization as a key for data warehouses
- Building a dimensional Module
- What are Dimensions?
 - Creating Dimension Tables
 - Degenerated Dimensions
 - Conformed Dimensions
- What are FACT tables?
 - Transaction FACT tables
 - Periodic Snapshot FACT tables
 - Accumulating Snapshot FACT tables
 - Fact Table Granularity – Defining the Grain
- What are Measures?
 - Denormalized tables Vs Normalized tables
 - Star Schema Vs Snowflakes Schema
 - Designing a Multidimensional Module
 - Examining the Business Processes of the Company
 - Prioritizing and building a stepped model
 - Building the Bus Matrix
 - Identifying the FACTs
 - Defining the Grain
 - Identifying the Dimensions
 - Examining real industry test cases
- Final Project – Modeling a real-life business process

פלטפורמת BI – ליבת הקורס

SSIS – SQL Server Integration Service

1. Introduction to ETL
 - What is Extract Transform Load software?
 - Understanding the ETL Architecture
 - Presenting alternative architectures
 - Exploring SSIS – SQL Server Integration Server and its Components
2. Building an ETL Solution using SSIS
 - Creating Packages
 - Working with data flows
 - Performing Basic Extracts
 - Loading data from a CSV file into SQL Server DB
 - Loading data from Multiple files into SQL Server DB
 - Loading data from a table into SQL Server DB
 - Loading data from multiple tables into SQL Server DB
3. Performing Incremental Load
 - Configuring a CDC (Change Data Capture)
 - Performing an Incremental Load using CDC
4. Performing Data Transformation
 - Union
 - Lookup
 - Sorting
 - Redirecting unmatched rows
 - Data Conversion
 - Derived Columns
 - Conditional Split
 - Data Verification
5. Error Handling and Debugging
 - Configuring DFT Error Output and Redirecting bad data
 - Loading error records into a table with specific error description (Script Task)
 - Error Handling by using Event Handlers
6. Control Flows
 - What is a control flow?
 - Working with Bulk Insert Task
 - Variables and Row Count
 - Implementing SCD using SSIS
7. Staging Area
 - The purpose of staging area
 - Using a staging area
 - Implementing solution using PT tables and Staging area
7. Enforcing Data Quality
8. ETL Design
 - Creating an ETL solution End-To-End
 - Administration and Monitoring
6. Package Deployment
 - a. Deploying a package using SSDT and executing it on SQL Server
 - b. Industry best practices

פלטפורמת BI – ליבת הקורס

Data Visualization Fundamentals

- לימוד עקרונות ה Data Visualization,
- הכרות בסיסית עם Tableau
- הכרות עם Power BI
- התחברות למקור נתונים
- שליפת נתונים ממקורות שונים (כמו אקסל, אתרי אינטרנט, SQL Server ועוד)
- ביצוע Data Cleansing
- הכרת שלשת המודולים של Power BI : Data view, Report view, Relationships view
- ביצוע טרנספורמציה של הנתונים
- בניית מודל הנתונים הפנימי עליו עובד Power BI המחובר למקור נתונים חיצוני
- עבודה עם ה Query Editor
- בניית עמודות חדשות על סמך המבנה הקיים
- יצירת Calculated Columns
- יצירת Measures
- בניית הקשרים הבין טבלאיים בתוך המודל הפנימי
- Normalization Vs Denormalization
- עבודה עם שפת DAX – פונקציות, פונקציות X, פונקציות לטיפול בתאריכים
- המחשה ויזואלית של הנתונים
- בניית Dashboards אינטראקטיביים
- בניית Reports המורכבים מכמה Dashboards המחוברים ביניהם ויצירת Story.

פלטפורמת BI – ליבת הקורס

Big Data

- מה זה בעצם Big Data?
- שימושים ויישומים מעשיים
- הקשיים והאתגרים הניצבים לפנינו בעולם ה Big Data
- סקירת ארכיטקטורות קלאסיות כגון Client – Server, Master -Slave
- Distributed File System והבנה מדוע הארכיטקטורות של "העולם הישן"
- אינן מסוגלות ליתן מענה לאתגרים שמציב עולם ה Big Data
- פריצת הדרך – Google File System
- Map & Reduce
- Hadoop Distributed File System
- Hadoop Echo Systems, שלשת השכבות מעל Hadoop
- Hive
- Pig
- Spark, הבנת פריצת הדרך החדשה של Spark
- מסדי נתונים מסוג NoSQL לעומת RDBMS
- Document Vs Table

Machine Learning

מבוא ל - Machine Learning

1. מבוא לתכנות

- מהו אלגוריתם?
- משתנים
- טיפוסים נתונים
- לולאות
- משפטי התנייה
- פונקציות
- תרשימי זרימה
- מערכים

2. מבוא לסטטיסטיקה

- מושגי יסוד בסטטיסטיקה
- חישובים סטטיסטיים והסתברותיים

Machine Learning

מימוש מודלים של שפת מכונה

1. תכנות Python

- הכרת סביבת העבודה
- כתיבת קוד ב-Python
- הכרת ספריות העזר המרכזיות
- כתיבת תוכנות ב-Python

2. מודלים סטטיסטיים של שפת מכונה ויישומם ב-Python

- Regression: Simple Linear Regression, Multiple Linear Regression, Polynomial Regression, SVR, Decision Tree Regression, Random Forest Regression
- Classification: Logistic Regression, K-NN, SVM, Kernel SVM, Naive Bayes, Decision Tree Classification, Random Forest Classification

3. נושאים מתקדמים בשפת מכונה

קורס זה ילווה בתרגילים מעשיים רבים ובסיומו פרויקט סיום.



המרכז הבינלאומי
ללימודי הייטק וחדשנות

* 6377

מתקדמים
לקריירה בהייטק



Microsoft Partner
Gold Learning



קמפוסים בפריסה ארצית:

באר שבע

רחוב האנגריה 77
פארק ההייטק

ירושלים

רחוב יפו 34

רחובות

רחוב אופנהיימר 5
פארק המדע

תל אביב

ראול ולנברג 36
קריית עתידים