

**FEDERAL TVET INSTITUTE (FTVETI)**

**Yeka subcity, Addis Ababa**

**Department of Information and Communication Technology**

**Big Data Analytics course catalog**

**Program:** M. Sc. in Information and Communication Technology. **I year II semester**

|  |
| --- |
| **1. Instructor Information** |
| **Name** | Dr. VASU PINNTI |
| **Office Location** | BLOCK-2, ROOM:204, ICT DEPARTMENT |
| **Phone Number** | +251936455539 |
| **E-mail** | vasupinninti@gmail.com |
| **2. Course Information** |
| **Course Name /Course Code** | Big Data Analytics |
| **Credit hours** | 3 Credit Hour , 2 Lecture ,3 Hour Lab |
| **3. Course Description** |
| **Aim**: After the end of this course the student will be able to acquire the skill and knowledge of Big data analytics techniques.**Description:**The main goal of this course is to help students learn, understand, and practice big data analytics and machine learning approaches, which include the study of modern computing big data technologies and scaling up machine learning techniques focusing on industry applications. Mainly the course objectives are: conceptualization and summarization of big data and machine learning, trivial data versus big data, big data computing technologies, and machine learning techniques.Students then engage in case study exercises in which small groups of students develop and present a big data concept for a specific real-world case. This includes practical exercises to familiarize students with the format of big data. The course is designed as a primer for anyone interested in attaining a basic understanding of what big data analysis entails.  |
| **4. Method of Instruction** |
| Class lectures /Demonstrations | In-class Tutorial (Evaluation) |
| Lab practice/ assignments | Individual/Group Assignment |
| mid/final semester examination | Group project |
| **5. Learning Outcomes** |
| **After the completion of this course the student will be able to:** |
| **5.1** | Students will demonstrate knowledge of big data analytics..  |
| **5.2** | understand about Big Data growth, the limitations of the existing solutions for Big Data problem, |
| **5.3** | Understand the various components of Hadoop architecture, how Hadoop solves the Big Data problem |
| **5.4** | Learn Hadoop Distributed File System (HDFS) make sense of how to function with them for limit and resource organization |
| **5.5** | Understand MapReduce and its qualities and retain advanced MapReduce thoughts |
| **5.6** | demonstrate the ability to think critically in making decisions based on data analytics |
| **5.7** | demonstrate the ability to use technical skills in predicative and prescriptive modeling to support business decision-making. |
| **6. Course Outline** |
| **Chapter** | Week | **Topics to be covered(Lecture hours)** | **Learning Outcomes** |  |
| **Chapter:1**Introduction to Big Data. | **1,2** | * What is Big Data?
* distributed file system vs Big Data Characteristic of Big Data( Five Vs)
* Big Data sources
* Challenges of Big Data
* How Hadoop solves the Big Data problem
* Tools used in Big Data
* Benefits of Big data
* Application of Big Data
 | **5.1, 5.2** |  |
| **Chapter: 2**Introduction to HADOOP. | **3-4** | * What is Hadoop
* A Brief History of Hadoop
* Big data analytics with Hadoop
* Hadoop EcoSystem projects
* Hadoop distributions
* Hadoop usecases
 | **5.3** |  |
| **Chapter: 3**Handling Big Data using Hadoop DFS | **5-7** | * HDFS-Overview and Goals
* HDFS architecture
* NameNode and DataNodes
* Data placement and Replication
* Robustness
* Accessibility
 | **5.4** |  |
| **Chapter: 4** Processing Big Data using map reduce | **8-11** | * What is map reduce and why it is so popular
* Map reduce big picture: map and reduce
* Map reduce process and terminology
* Map reduce components failures and recovery
 | 5.5 |  |
| **Chapter 5**machine learning with Big data | **11-13** | * Introduction to machine learning
* Supervise learning
* Unsupervised learning
 | 5.6., 5.7 |  |
|  |  |  |  |  |
| **8. Required Text and Reference** |
| **Text Book** | * ***Big Data: A Revolution That Will Transform How We Live, Work, and Think*** by Kenneth Cukier and Viktor Mayer-Schönberger

Big Data For Dummies by [Judith S. Hurwitz](https://www.amazon.com/s/ref%3Ddp_byline_sr_book_1?ie=UTF8&field-author=Judith+S.+Hurwitz&text=Judith+S.+Hurwitz&sort=relevancerank&search-alias=books) , [Alan Nugent](https://www.amazon.com/Alan-Nugent/e/B0034PR9UC/ref%3Ddp_byline_cont_book_2)  , [Fern Halper](https://www.amazon.com/s/ref%3Ddp_byline_sr_book_3?ie=UTF8&field-author=Fern+Halper&text=Fern+Halper&sort=relevancerank&search-alias=books) , [Marcia Kauf](https://www.amazon.com/Marcia-Kaufman/e/B00B29M9K2/ref%3Ddp_byline_cont_book_4)man* Hadoop: The Definitive Guide, Tom White, O’Reilly
 |
| **Reference Books** | * Matthew J. Salganik. (2017). Bit by Bit: Social Research in the Digital Age. Princeton University Press.
* Cathy O’Neil. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Penguin Books.
* Rob Kitchin. (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. SAGE Publications.
 |
| **9. GRADE Assesment** |
| **Type** | **Weight** | **submitting date** |  |
| **Mid semester Exam** | 20% | 8th week of the semester |  |
| **Final semester Exam** | 40% | 15th /16th week of the semester |  |
| **Lab Practice/observation** | 20% | 2 or 3 consecutive weeks |  |
| **Group assignment / Project** | 20% | the last two weeks of the semester end |  |
| **Attendance:** You are obliged to attend at least 80% of the lecture classes and 100% attendance during practical work sessions to sit for final exam |
| **10. Academic Honesty** |
| In all cases i.e. in performing assignments, laboratory works, project works and examinations, copying from others and using others’ work as own is considered to be cheating and cheating is forbidden by the law of the academic principles and regulation of the institute. Failure to do so will lead to take disciplinary action starting from canceling the results of the assignments; project works, laboratory activities and examination up to the dismissal of the institute in consultation with the concerned body. |
| **11. Submitting Date** |
| All assignments, project works and laboratory reports should be submitted to the instructor or laboratory assistance according to the timetable provided. All assignments, project works and laboratory results will be invalid if they are not submitted on time and will be reported to the concerned body as miss conduct of the student. |
| **12. Classroom Behavior** |
| Classroom discipline is primary for healthy teaching learning process. Therefore, it is the responsibility of the class to avoid disturbing behaviors and activities that competes the attention of the class and the instructor. Switching off the cell phone is vital in the class. Failure to do so will lead to take disciplinary measure. |
| **13. Approval (Affidavit)** |
| **Name** | **Signature** | **Date** |
| **Instructor: Dr vasu pinnti** |  |  |
| **Section Head: Prof Ravindra Babu** |  |  |
| **Department Head: Ato Sisay** |  |  |