**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**

**WORK INTEGRATED LEARNING PROGRAMMES**

**Digital**

Part A: Content Design

|  |  |
| --- | --- |
| **Course Title** | AI and ML Techniques for Cyber Security |
| **Course No(s)** |  |
| **Credit Units** | 5 |
| **Credit Model** | 1 - 0.5 - 1.5.  1 unit for class room hours, 0.5 unit for Tutorial, 1.5 units for  Student preparation. 1 unit = 32 hours |
| **Content Authors** | Ms. Seetha Parameswaran |
| **Version** | 1.0 |
| **Date** | January 20th, 2021 |

**Course Objectives**

|  |  |
| --- | --- |
| **No** | Course Objective |
| **CO1** | To introduce the basic concepts and techniques of Artificial Intelligence and Machine learning. |
| **CO2** | To apply Machine Learning techniques to Cyber Security. |
| **CO3** | To analyse the performance of applied Machine Learning models. |

**Text Book(s)**

|  |  |
| --- | --- |
| T1 | Data Mining and Machine Learning in Cybersecurity, Sumeet Dua and Xian Du, CRC Press, 2011 |
| T2 | Artificial Intelligence & Cybersecurity For Dummies®, IBM Limited Edition, 2018 by John Wiley & Sons, Inc |
| T3 | Introduction to artificial intelligence for security professionals, The Cylance Data Science Team. Irvine, CA : The Cylance Press, 2017 |
| T4 | Hands-On Machine Learning for Cybersecurity, Soma Halder and Sinan Ozdemir, Packt publishing |

**Reference Book(s) & other resources**

|  |  |
| --- | --- |
| R1 | William Stallings & Lawrie Brown, Computer Security: Principles and Practice, 4th Edition, Pearson, 2018 |
| R2 | Tom M. Mitchell, Machine Learning, The McGraw-Hill Companies, Inc. Indian Edition 1997 |
| R3 | Christopher M. Bishop, Pattern Recognition & Machine Learning, Springer, 2006 |

**Content Structure**

1. Introduction to Cyber Security ( 2 hrs)
   1. Introduction to cyber security, cyber security challenges, types of attacks
   2. Techniques for addressing the challenges, Block chain, Cryptographic techniques, Access control techniques, Mathematical approach, Soft computing techniques, AI and ML techniques.
   3. Review of Cybersecurity Solutions
   4. Objective of course AI-ML techniques for cyber security.
2. Introduction to Artificial Intelligence ( 2 hrs)
   1. Assisted intelligence, Augmented intelligence, Autonomous intelligence
   2. Predictive analytics
   3. Introduction to Cognitive computing
   4. Case Study: Cognitive security with IBM Watson
3. Basics for Machine Learning ( 4 hrs)
   1. Types of data: structured and unstructured data; labeled and unlabeled data
   2. Types of learning: supervised, unsupervised, reinforcement
   3. Data selection and sampling
   4. Feature extraction
   5. Feature encoding, vectorization, normalization
   6. Issues: Overfitting, Under fitting, Class Imbalance
   7. Evaluation metrics: precision, recall and F1-score
   8. Overview of Machine learning algorithms
4. Supervised Learning for Misuse/Signature Detection ( 2 hrs)
   1. Rule-Based Signature Analysis
   2. Decision Tree for Misuse Detection
   3. Artificial Neural Network for Misuse Detection
   4. Support Vector Machine for Misuse Detection
5. Machine Learning for Anomaly Detection ( 2 hrs)
   1. Association Rules in Audit Data for Anomaly Detection
   2. ANN Approach for Anomaly Detection
   3. Random Forest Approach for Anomaly Detection
   4. Clustering for Anomaly Detection
   5. Deep learning techniques for Anomaly Detection
6. Machine Learning for Malware detection and classification ( 4 hrs)
   1. DBSCAN clustering algorithm for Malware detection
   2. Random Forest for Malware Classification
   3. Malware detection and classification
   4. Android malware detection
   5. Deep learning techniques for Malware Classification
7. Network Intrusion detection and classification ( 4 hrs)
   1. Machine Learning in Hybrid Intrusion Detection Systems
   2. Anomaly–Misuse Sequence Detection System using Association rules
   3. Misuse–Anomaly Sequence Detection System using Random forest
   4. Intrusion detection
   5. AdaBoost-Based Machine Learning for Network Intrusion Detection
   6. Deep learning techniques for Intrusion Detection Systems
8. Detection and categorization of domain names generated by Domain name generation algorithms ( 2 hrs)
   1. Hidden Markov model for classification
   2. Deep learning technique helps in discrimination of DGA domains and non-DGA domains
9. Profiling Network Traffic ( 2 hrs)
   1. Machine Learning for Profiling Network Traffic
   2. Using clustering algorithms

1. Adversarial Machine Learning for Malware detection( 2 hrs)

**Learning Outcomes:**

|  |  |
| --- | --- |
| No | Learning Outcomes |
| LO1 |  |
| LO2 |  |
| LO3 |  |

**Part B: Learning Plan**

|  |  |
| --- | --- |
| **Academic Term** |  |
| **Course Title** | AI and ML Techniques in Cyber Security |
| **Course No** |  |
| **Lead Instructor** |  |

|  |  |  |
| --- | --- | --- |
| Session No. | Topic Title | Study / HW Resource Reference |
| 1 | * Introduction to Cyber Security   + Introduction to cyber security, cyber security challenges, types of attacks   + Techniques for addressing the challenges, Block chain, Cryptographic techniques, Access control techniques, Mathematical approach, Soft computing techniques, AI and ML techniques.   + Review of Cybersecurity Solutions   + Objective of course AI-ML techniques for cyber security. | R1  Web references |
| 2 | * Introduction to Artificial Intelligence   + Assisted intelligence, Augmented intelligence, Autonomous intelligence   + Predictive analytics   + Introduction to Cognitive computing   + Case Study: Cognitive security with IBM Watson | T2: Ch2, Ch5 Web references |
| 3 | * Basics for Machine Learning   + Types of data: structured and unstructured data; labeled and unlabeled data   + Types of learning: supervised, unsupervised, reinforcement   + Data selection and sampling | T4: Ch1 |
| 4 | * Basics for Machine Learning   + Feature extraction   + Feature encoding, vectorization, normalization   + Issues: Overfitting, Under fitting, Class Imbalance   + Evaluation metrics: precision, recall and F1-score   + Overview of Machine learning algorithms | T4: Ch1 |
| 5 | * Supervised Learning for Misuse/Signature Detection   + Signature Detection   + Rule-Based Signature Analysis   + Decision Tree for Misuse Detection   + Artificial Neural Network for Misuse Detection   + Support Vector Machine for Misuse Detection | T1: Ch3 |
| 6 | * Machine Learning for Anomaly Detection   + Anomaly Detection   + Association Rules in Audit Data for Anomaly Detection   + ANN Approach for Anomaly Detection   + Random Forest Approach for Anomaly Detection   + Clustering for Anomaly Detection   + Deep learning techniques for Anomaly Detection | T1: Ch 4  Web references |
| 8 | Review of Sessions 1 to 7 | Books, Slide deck |
| 9 | * Machine Learning for Malware detection and classification   + Malware detection and classification   + DBSCAN clustering algorithm for Malware detection   + Random Forest for Malware Classification | Web references |
| 10 | * Malware detection and classification   + Android malware detection   + Deep learning techniques for Malware Classification | Web references |
| 11 | * Intrusion detection   + Machine Learning in Hybrid Intrusion Detection Systems   + Anomaly–Misuse Sequence Detection System using Association rules   + Misuse–Anomaly Sequence Detection System using Random forest | T1: Ch5 |
| 12 | * Intrusion detection   + AdaBoost-Based Machine Learning for Network Intrusion Detection   + Deep learning techniques for Intrusion Detection Systems | T1: Ch5 |
| 13 | * Detection and categorization of domain names generated by Domain name generation algorithms   + Hidden Markov model for classification   + Deep learning technique helps in discrimination of DGA domains and non-DGA domains | Web references |
| 14 | * Profiling Network Traffic   + Machine Learning for Profiling Network Traffic   + Using clustering algorithms | T1: Ch7 |
| 15 | * Adversarial Machine Learning for Malware detection | Web references |
| 16 | Review of Sessions 9 to 15 | Books, Slide deck |

**Detailed Plan for Lab work**

|  |  |  |  |
| --- | --- | --- | --- |
| **Lab No.** | **Lab Objective** | **Lab Sheet Access URL** | **Module Reference** |
| 1 | Introduction to the tools and installations |  | - |
| 2 | Misuse/Signature Detection |  | 4 |
| 3 | Spam Classification |  | - |
| 4 | Anomaly Detection |  | 5 |
| 5 | Malware detection |  | 6 |
| 6 | Malware classification |  | 6 |
| 7 | Anomaly–Misuse Sequence Detection |  | 7 |
| 8 | Profiling Network Traffic |  | 9 |

**Evaluation Scheme**:

Legend: EC = Evaluation Component; AN = After Noon Session; FN = Fore Noon Session

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No | Name | Type | Duration | Weight | Day, Date, Session, Time |
| EC-1 | Quizzes | Online |  | 10% |  |
|  | Assignments | Take Home |  | 20% |  |
| EC-2 | Mid-Semester Test | Closed Book | 1.5 Hrs | 30% |  |
| EC-3 | Comprehensive Exam | Open Book | 2.5 Hrs | 40% |  |

**Note:**

Syllabus for Mid-Semester Test (Closed Book): Topics in Session Nos. 1 to 8

Syllabus for Comprehensive Exam (Open Book): All topics (Session Nos. 1 to 16)

**Important links and information:**

Elearn portal: [https://elearn.bits-pilani.ac.in](https://elearn.bits-pilani.ac.in/) or Canvas

Students are expected to visit the Elearn portal on a regular basis and stay up to date with the latest announcements and deadlines.

Contact sessions: Students should attend the online lectures as per the schedule provided on the Elearn portal.

Evaluation Guidelines:

1. EC-1 consists of two Quizzes. Students will attempt them through the course pages on the Elearn portal. Announcements will be made on the portal, in a timely manner.
2. EC-1 consists of either one or two Assignments. Students will attempt them through the course pages on the Elearn portal. Announcements will be made on the portal, in a timely manner.
3. For Closed Book tests: No books or reference material of any kind will be permitted.
4. For Open Book exams: Use of books and any printed / written reference material (filed or bound) is permitted. However, loose sheets of paper will not be allowed. Use of calculators is permitted in all exams. Laptops/Mobiles of any kind are not allowed. Exchange of any material is not allowed.
5. If a student is unable to appear for the Regular Test/Exam due to genuine exigencies, the student should follow the procedure to apply for the Make-Up Test/Exam which will be made available on the Elearn portal. The Make-Up Test/Exam will be conducted only at selected exam centres on the dates to be announced later.

It shall be the responsibility of the individual student to be regular in maintaining the self-study schedule as given in the course hand-out, attend the online lectures, and take all the prescribed evaluation components such as Assignment/Quiz, Mid-Semester Test and Comprehensive Exam according to the evaluation scheme provided in the hand-out.