Harper Emerging Technology BILT: Cybersecurity BILT 6.20.2024

Cyber AI Knowledge Units (Kus)

This includes KUs focusing on ‘Security of AI’ and ‘AI in Cybersecurity’.

Knowledge Unit Name	AI Overview
Topics: ● AI principles ● AI terminology and definitions ● AI history ● What is AI and AI components ● AI careers ● Security of AI ● Human behavior and performance ● Future of AI in cybersecurity ● Basics in security in privacy ● Real world applications of AI ● AI applications in general and in cybersecurity ● AI examples (successes and failures) ● AI applications use cases ● Common AI tools and how to evaluate them ● Emerging applications	

Knowledge Unit Name	ML and AI Fundamentals for Cybersecurity

Topics:

- Machine learning, ML Models and Theory
- Natural Language Processing (NLP)
- Generative AI
- Prompt engineering
- Explainable AI
- Defining neural networks
- Training / test / validation data splits
- Parameter tuning
- Overfitting and mitigation
- Performance metrics
- Anomaly detection
- Supervised and unsupervised
- Machine learning, deep learning, neural networks
- Fundamentals of deep neural networks (K-nearest neighbors (KNN), Convolutional Neural Network (CNN), Long-Short-Term-Memory (LSTM), Graph Neural Network (GNN), Recurrent Neural Network (RNN), Neural Networks (NN), deep learning techniques, autoencoders, transformer models, generative AI)
- ML Fundamentals (math foundations, math models, ML algorithms, ML tuning, AI tools)
- AI concepts (AI lifecycle, AI case studies, prompt engineering, data, history, general issues in AI)

Knowledge Unit Name	Legal and Ethical Issues of AI

KU Learning Outcomes (student will be able to):

1. Assess AI trustworthiness
2. Describe trust vs trustworthiness
3. Define trustworthiness requirements for an AI system

Topics:

- AI Ethics (Equitable AI, Fairness, Biases, Misinformation, Monetary validation)
- Equitable AI (Metrics, Methods)
- AI Ethics (How to quantify / How to protect)
- Bias and fairness of training data
- Trust and related attributes / metrics
- Explainability tools and visualizations
- Calibrated trust
- Certification authorities
- Trustworthy AI taxonomies
- AI uses, benefits, risks
- Ethical concerns

Knowledge Unit Name	Data Protection and Privacy

Topics:

- Uncertainty management
- AI Security algorithm techniques
- AI privacy protection methods metrics
- AI privacy risk assessments - model infrastructure ● Privacy
- Security (network security, database administration)
- Governance
- Policy (HIPAA, regulations, retention)
- Privacy (noisy data, encryption)
- Documentation
- Human Subjects (training)

Knowledge Unit Name	Data Sources

Topics:

- Identification
- Collection strategies
- Data Characteristics
    - Data modality
    - Class / balance issues
    - Data generation
    - Formats and standards
    - Data providence
    - Volume, variety, and velocity

Knowledge Unit Name	Math Foundations

Topics:

- Linear combination and matrix multiplication
- Probability and statistics
- Definition of probability distributions
- Definitions of derivatives, partial derivatives
- Mathematical optimization
- LP norms and vectors

<table>
<tr><td>Knowledge Unit Name</td><td>AI Algorithms</td></tr>
<tr><td colspan="2">Topics:
• AI programming
• Future engineering
• Rule based systems
• Programming design
• Linear and logistics regression
• Random forests
• Planning search-based planning
• Support vector machines
• Decision trees
• Naive bayes
• Knowledge base and representation
• Clustering algorithms</td></tr>
</table>

<table>
<tr><td>Knowledge Unit Name</td><td>Exploratory Data Analysis</td></tr>
<tr><td colspan="2">Topics:
• Summary statistics
• Data visualization
• Feedback loops</td></tr>
</table>

<table>
<tr><td>Knowledge Unit Name</td><td>Problem Discovery</td></tr>
<tr><td colspan="2">Topics:
• Problem specification ○ Damage assessment and reduction
• Domain specific requirements
• Formulating cybersecurity tasks of ML problems (identify objective, data, performance targets)
• Communication to non-technical audience
• Identifying problems that can be addressed via ML
• Risk and societal impact of AI in context of applications
• Supply chain (minimize cyber risk)
• Domain-specific expertise for threat detection, identification, and mitigation</td></tr>
</table>

Knowledge Unit Name	Data Processing and Curation

Topics:

- Data Cleaning (Outlier detection, deduplication, normalization, data type, generation, augmentation, imputation
- Representation (vectorization, feature engineering / extraction, graph structure, embedded)
- Feature engineering for Machine Learning for Security (ML4Sec) (categorical vs numerical, embeddings, feature selection / projection
- Data availability ○ Data ownership ○ Data Integrity
- Sources of data for ML4Cyber (issues with privacy, imbalance, incompleteness, distributional shift; modalities of data)
- Intelligence Analysis / OSZN

Knowledge Unit Name	Model Selection and Specification

Topics:

- Multimodal system for security
- Optimization of COP (Tools, Operators, Third-party)
- ML Training (model selection / design, training performance monitoring, benchmarking and performance education, class imbalance / regulation)
- AI blindspot ○ Model implementation
- Model maintenance
- Training objectives in adversarial nations universities
- Familiarization with vulnerability of AI system as a hybrid SW / HW system ○ Limitation of assumption of AI tools

Knowledge Unit Name	Model Evaluations

Topics:

- Adversarial examples for the problem
- Model Resilience (adversarial evaluation / red teaming, performance monitoring, tracking distributional shift, recalibration, versioning)
- Evaluation metrics
- Understand AI models and capabilities
- Defense mechanisms ○ Robustness metrics
- AI model performance optimization for false alarm reduction ○ AI explain ability / output interpretation

Knowledge Unit Name	Security Assessment and Evaluation

KU Learning Outcomes (student will be able to):

1. Perform threat modeling of AI systems
2. Monitor and detect threats of AI systems
3. Mitigate ML and AI threats
    a. Understand AI forensics
    b. Perform incident response and recovery
4. Continuous assessment
5. Identify attack surfaces and threat model
6. Evaluate and select benchmarking framework
7. Perform security evaluation and produce report

Topics:

- Threat modeling
- Monitoring and threat detection
- AI forensics
- Benchmarking and education
- Taxonomy of attacks and defense
- Red teaming tools and frameworks
- Monitoring and detection
- AI frameworks

Knowledge Unit Name	Risk Management of AI

Topics:

- Metrics qualification
- Continuous monitoring
- Risk analysis: data
- AI risk qualifications / metrics
- Risk assessment
- Risk management
- Risk communication
- Data engineering
- Human in the loop / top
- AI risk mitigation
- Mitigation planning (accept / mitigate)
- Risk management failures
- Continuity of operations
- Scenario planning or related AI risks
- Algorithms
- Models
- Systems / Applications

Knowledge Unit Name

Databases and Infrastructure

Topics:

- Databases (Relational, NoSQL, Vector, Hybrid, Graph, Anonymization)
- Infrastructure (Cloud, Networking, Storage, Data Lake)
- Centralized and Distributed (cross cutting)

Knowledge Unit Name

System Deployment and Operation

Topics:

- Damage investigation
- Machine Learning Operations (MLOPS) (data pipeline, deployment infrastructure engineering (cloud vs on-prem vs edge)
- Deployment and integration of Cyber AI applications
- AI system vulnerability (Dynamic)

Knowledge Unit Name

Defensive Applications of AI

Topics:

- Cyber threat intelligence
- AI-based NIDS
- Network security for deep learning
- AI for web security
- Automated system security compliance scanner, notification, remediation, at risk reporting
- Psychological modeling of an insider threat
- Reinforcement learning and robotics
- AI for blue team
- AI forensics
- AI-based vulnerability defense / repair
- Malware analysis
- Privacy preserving AI

Knowledge Unit Name

Offensive Applications of AI

Topics:

- AI for malware curation
- AI red-teaming tool for training
- AI for offensive security
- AI for social engineering
- AI for social media (deepfake, false news, hate speech)
- AI in misinformation
- AI for cyber deception

Knowledge Unit Name	Adversarial Learning

KU Learning Outcomes (student will be able to):

1. Craft an advanced attack using a given a model and dataset
2. Refrain the model and reassess robustness
3. Understand taxonomy of attacks and defense
    a. Conduct a data-poisoning attack
    b. Detect and prevent a data-poisoning attack

Topics:

- Taxonomy of attacks and defenses
- Formulation of attacker objectives
- Formulation of defender objectives
- Mapping objectives to algorithms
- Measuring robustness correctly

Knowledge Unit Name	Security and Governance of AI

KU Learning Outcomes (student will be able to):

1. Identify vulnerabilities in connections between AI components
2. Connect AI models and mitigate risk from vulnerabilities

Topics:

- Proactive Defense
- Reactive Defense
- Mitigation Operations (How to execute mitigation plans)
- Privacy
- AI applications (levels of severity, types)
- AI blind spot
- Physical systems controls (electrical grid / access control)
- Keep AI safe from various dangers (physical and cyber threats)
- Potential AI security risks
- Cybersecurity
- Perform direct prompt injection attacks and defense
- Perform indirect prompt injection attacks and defense
- Understand generative AI attacker goals and objectives
- Generate and detect unsafe synthetic data
- Security of federated learning systems
- Security of swarm systems
- Graceful degradation when components lose connectivity
- Security of connected AI models

Knowledge Unit Name

Lifecycle Management of AI

KU Learning Outcomes (student will be able to):

1. Identify risk in AI lifecycle
2. Differentiate threats at different stages of AI lifecycle
3. Implement mitigation techniques

Topics:

- Terminate or retire AI systems that don’t meet the organization’s value or standards
- Validation of models
- Certification of AI System
- Ingrain AI into your organization culture before deployment
- Stakeholder engagement and communication
- Data science security and data flow security
- Human-model interaction
- Server authentication of component integration (Third-party?)

Knowledge Unit Name

Regulation and Governance of AI Risks

Topics:

- AI policies and compliance
- Policy and regulation compliance
- Manage, evaluate, and hold AI accountable
- Onboard AI as your organizations builds new employees and Third-party vendors
- Build integrity into your organization’s AI from the design stage
- Intelligence Community (IC) risk framework
    - Global Competition
    - IC use of AI
    - No US person data