CERIAS

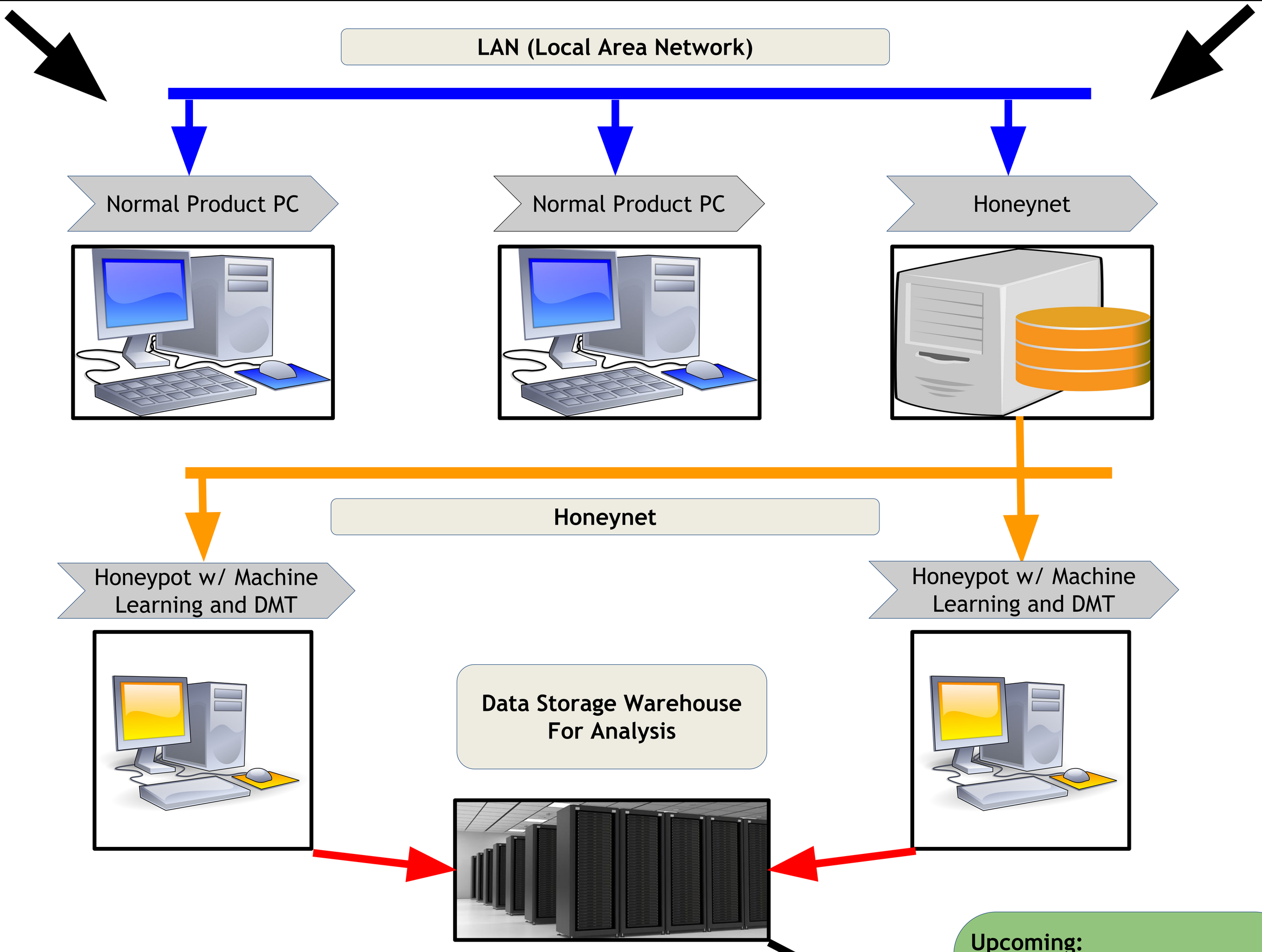
The Center for Education and Research in Information Assurance and Security

Exploring Deception with Machine Learning and Data Mining

Presenter (Independent Undergraduate): Wesley LaFleur

Problem Statement:

- Only **2-3%**[1] of companies share breach info, which could hold **useful information that could be beneficial for threat detection/prevention companies**
- Current deception software allows information to be **observed**, but **not stored**[2]
- **General understanding** of the **techniques** used by data mining tools is **limited for cyber security**, due to what **they were originally designed for**[3]



1. Bahmani, B. (2014). *CS259D: Data Mining for CyberSecurity* (1st ed., p. 5). Stanford University. Retrieved from http://web.stanford.edu/class/cs259d/lectures/Session13.pdf
2. Richardson, R. (2015). *Security startups tackle the art of deception techniques*. *SearchSecurity*. Retrieved 3 April 2016, from http://searchsecurity.techtarget.com/opinion/Security-startups-tackle-the-art-of-deception-techniques
3. McKnight, W. (2007). *Data mining tools*. *SearchBusinessAnalytics*. Retrieved 3 April 2016, from http://searchbusinessanalytics.techtarget.com/answer/Data-mining-tools-Advantages-and-disadvantages-of-implementation

Upcoming:

- Further investigation into costs lost/gained, reliability lost/gained, time lost/gained, etc.
- Tests using deceptive software with machine learning applied to it.

PURDUE UNIVERSITY