Capacity Building: Integrating Data Science into Cybersecurity Curriculum

Edoardo Serra, Francesca Spezzano (Boise State University) and Dianxiang Xu (University of Missouri-Kansas City)
edoardoserra@boisestate.edu

BOISE STATE UNIVERSITY

Goal:

Develop innovative curriculum materials on security data science by integrating data science workflow, security problems, the adversary's perspective of security, and inquiry-based learning into hands-on practices.

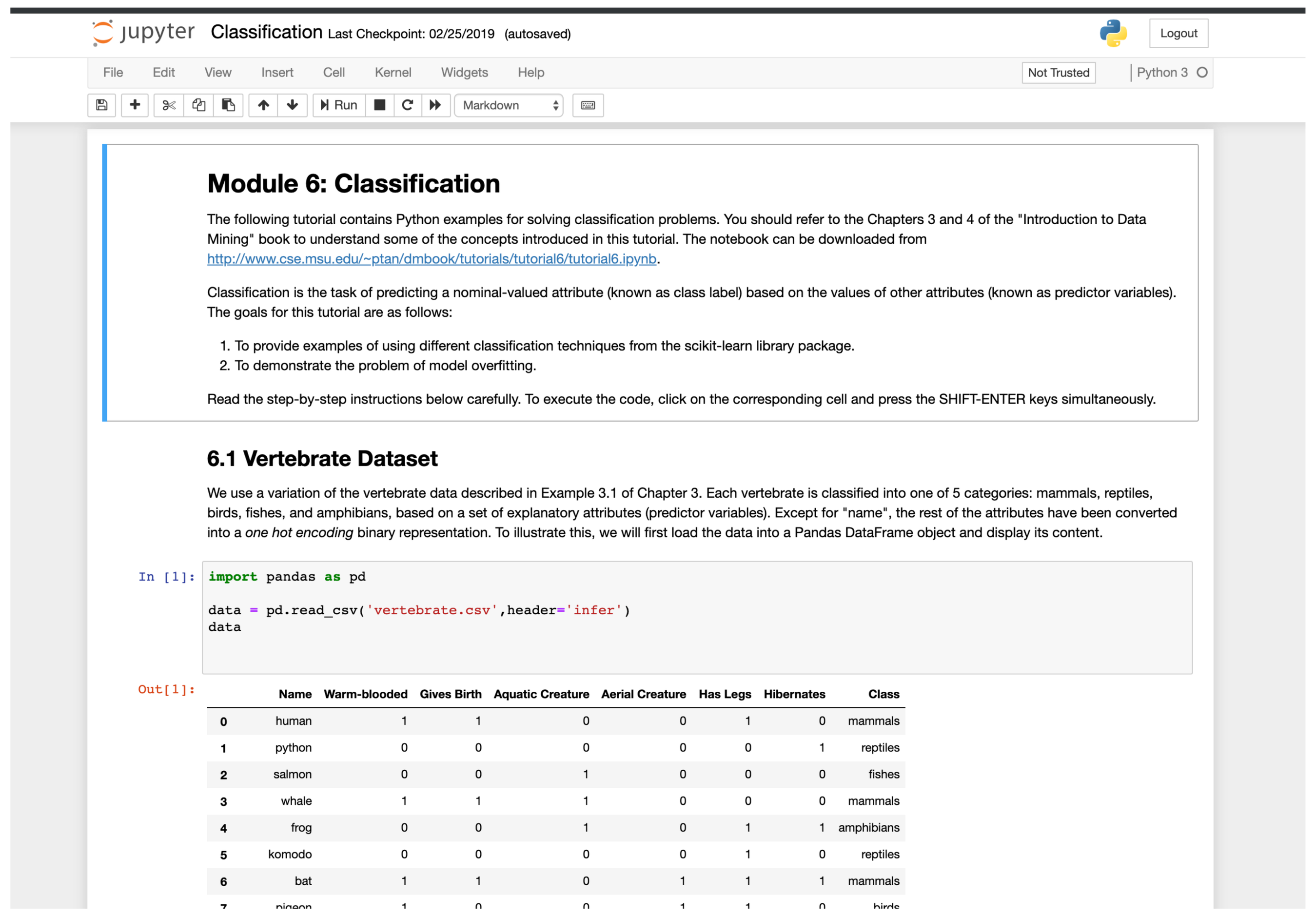
Objectives:

- **Develop innovative hands-on curriculum materials** (e.g., lecture notes and lab manuals) on security data science that features inquiry-based learning and the adversary's perspective of security. The curriculum material covers the following topics: malicious user detection, intrusion detection, malware detection, and vulnerability prediction.
- **Develop an open collaborative repository for the security data science community**. This repository hosts the proposed curriculum material and serves the community as a collaborative environment for educators to exchange ideas, contribute new content and share resources.
- **Hold national faculty development workshops on security data science**. These workshops bring together various stakeholders such as educators, industry, practitioners to discuss needs and progress in security data science and how to integrate those in specific courses.

Curriculum Material:

Python notebook demonstrating how to detect opinion spammers in Yelp. We used a popular Yelp restaurant review dataset. Spammers, in our case, are users who wrote at least one filtered review. In the notebook, we start by showing how to load and clean the dataset and printing dataset statistics. Next, we initiate a discussion with the students where we use inquiry-based learning and stimulate students to think from the adversary perspective. Here we show some sample questions:

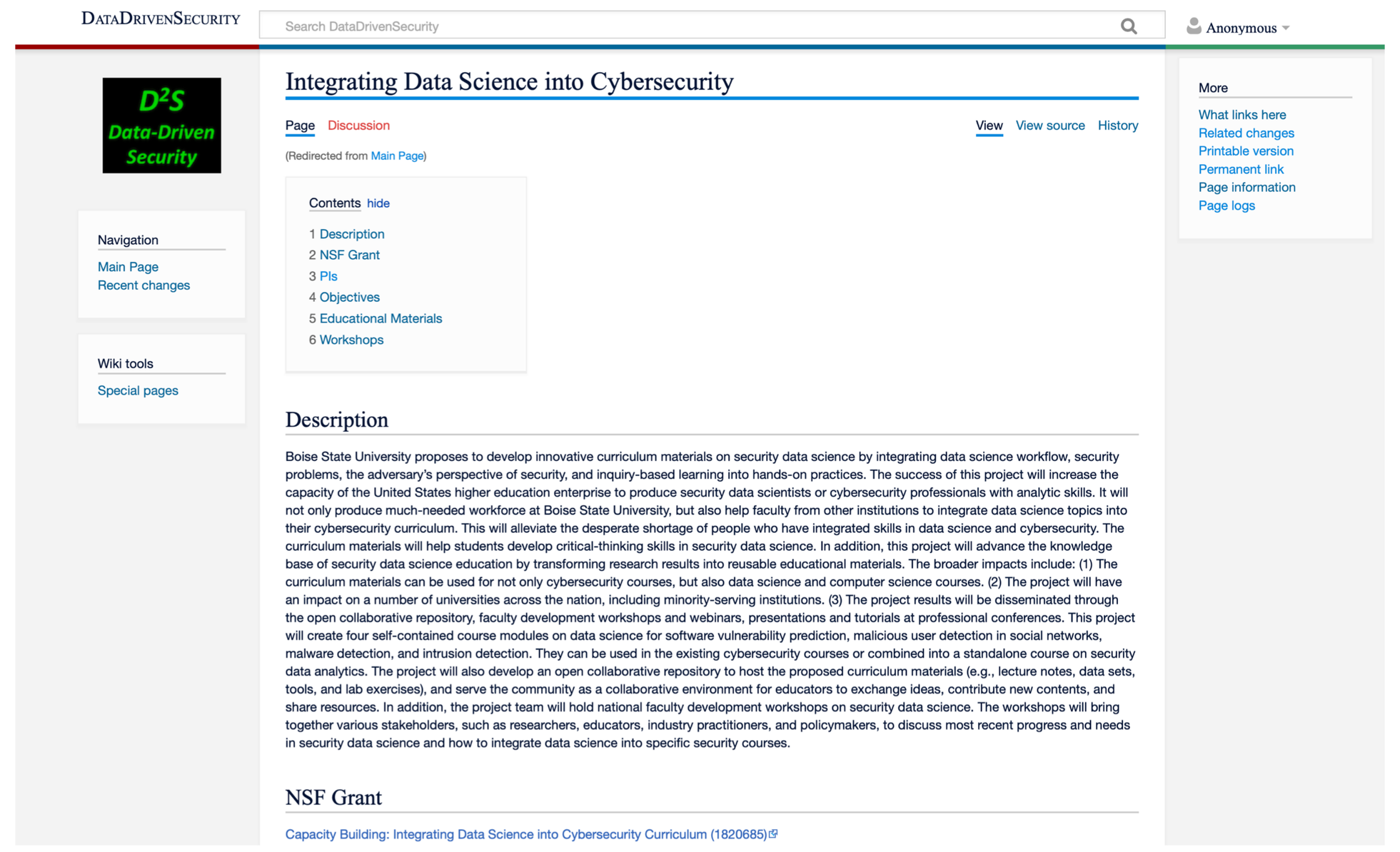
- "think of two or three features that can be used to detect opinion spammers"
- "think of how these features can describe the behavior of a spammer"
- "what can be done to improve the classification results and how?"



Repository:

We created the collaborative repository. The repository allows people to register and create their own content and share it with the community. We already populated the repository with basic information about the grant, the educational material that we have already developed and the 2019 workshop details and pictures about the event.

Repository Website: https://datadrivensecurity.boisestate.edu



Workshop 2019

We hosted the first 2-day faculty development workshop at Boise State this year in June. We had 30 participants coming from different Universities located across the nation.

External Evaluation: *The project has made significant progress towards meeting stated goals and objectives. Those engaged with project activities, and interested or associated with data science and/or cyber security, are recognizably diverse as it relates to areas of expertise, working environment and employment title (including both private and public sectors) and geographic area. Developed materials show a significant effort to integrate Data Science and Cyber Security. Regarding the workshop, pre/post questionnaires given to participants indicate an effective selection of speakers and the delivery of well-received day of instruction. In addition, feedback offered by respondents show high levels of satisfaction as well as opportunities for additional content for future events and available online materials.*

2019 Workshop Program

Day 1

- Project introduction
- Industry presentations
- Faculty presentations
- Government agencies presentations
- Educational material presentation and open discussion

Day2

- Participants networking for possible collaboration