Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Viktor Vašátko
Název práce	Applications of Artificial Intelligence in IT Security
Rok odevzdání	2017
Studijní program	Informatika **Studijní obor** Diskrétní modely a algoritmy

Autor posudku	Roman Neruda **Role** Oponent
Pracoviště	ÚI AV ČR, v.v.i.

Text posudku:

Práce Viktora Vašátka se zabývá aplikací metod strojového učení pro klasifikaci útoků na počítačové systémy. Cílem práce bylo vytvoření kompletního workflow strojového učení, které obsahuje metody předzpracování, vizualizace, redukce dimenze, učení modelu a extrakci pravidel z něj. Autor používá standardní reálná benchmarková data CSE-CICIDS2018 pro vývoj a testování svého modelu. Výsledkem práce je tak celá metodika tvorby modelu a zároveň sada klasifikačních pravidel, kterou lze potenciálně použít ve standardních pravidlových systémech detekce útoků.

Vlastní text práce je členěn do osmi kapitol a obsahuje rozsáhlou přílohu s podrobným popisem statistických vlastností dat a výsledky experimentů. Součástí práce je také kód v jazyce python implementující popsanou funkcionalitu za použití knihoven scikit-learn a skope-rules.
V úvodních nečíslovaných kapitolách autor stručně uvádí do problematiky bezpečnosti systémů a využití strojového učení, seznamuje s relevantními předchozími pracemi v tomto oboru a představuje strukturu celé práce. V kapitole 1 je popsána geneze vzniku a struktura datové sady CSE-CICIDS2018. Jde o velká data (16 milionů instancí, 80 příznaků) definující klasifikační úlohu s patnácti třídami. Četnosti tříd jsou velmi nevyvážené, což činí klasifikaci ještě obtížnější. Kapitola 2 je úvodem do strojového učení. V textu se popisují základní koncepty od typů úloh a ztrátových funkcí po konkrétní modely, které jsou využity v autorově vlastním řešení.
Toto řešení je představeno v kapitolách 3 a 4, které se týkají učení modelu a extrakce pravidel. Rozsáhlá třetí kapitola se věnuje transformaci dat (3.1), vizualizaci (3.2), využití metod redukce dimenze (3.3) a učení různých modelů (3.4). Ve čtvrté kapitole autor popisuje způsob extrakce pravidel z ansámblu rozhodovacích stromů a porovnává získané výsledky s původním modelem. Celá práce je shrnutá v závěrečné kapitole, kde autor komentuje dosažené výsledky a naznačuje možnosti další práce.

Z přínosů práce bych vyzdvihl následující:

1. Autor se problému strojového učení nad konkrétní úlohou věnuje v celém kontextu a popíše kompletní workflow od získání dat po použití extrahovaných pravidel. Každá část zpracování dat je prozkoumána co do použití různých metod a pečlivě statisticky zpracována. Dobře je popsána procedura extrakce pravidel a jejich použití v klasifikaci. Tím vznikl praktický postup využitelný v používaných bezpečnostních IT řešeních.

2. Balík skriptů v jazyce python implementuje navržené metody a umožňuje jejich snadné využití v praxi. Kvalita programového řešení je velmi dobrá, autor přizpůsobil metody ze standardních knihoven jejich použití pro konkrétní problém spadající bezesporu do kategorie big data. Materiál obsažený v příloze je dobrým základem k reprodukci autorových experimentů a pokračování ve vývoji.

Následuje několik poznámek a otázek k obhajobě:

1. Práce je psána srozumitelně a logicky, obsahuje jen malé množství drobných chyb v angličtině a typografických chyb.

2. Co považujete za největší vlastní přínos práce?

3. Jaké je srovnání vašich výsledků s jinými pracemi, zvlášť existují li modely založené na hlubokém učení?

4. V práci zmiňujete, že nedošlo k zamýšlenému využití pravidel v reálném nasazení. Můžete rozvést svou představu o způsobu nasazení vašeho řešení v praxi?

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhujete na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum 1. 7. 2020 **Podpis**