

Oponentní posudek bakalářské práce

Diplomant: Jaromír ŠÍDA

Téma: Příspěvek nasazení dataminingu k bezpečnosti ve společnosti

Cílem bakalářské práce bylo seznámit se s metodami dataminingu a vytvořit modelový případ v nástroji PSW Modeler. Pro tyto účely byly zvolené zajímavé vstupní data mapující kriminální činnost na určitém území, za konkrétní období. Dále měl bakalář naprogramovat funkční aplikaci, která demonstruje pochopení vybraného algoritmu. Pro účely výuky měl své poznatky bakalář shrnout do e-learningového kurzu.

Bakalář začal svou práci vysvětlením základních pojmů a technik dataminingu. Podrobněji se zaměřil na hierarchické i nehierarchické metody shlukové analýzy. Tyto metody umožňují řešit deskriptivní případy dataminingu, jakým je taky zvolená modelová úloha. Kapitoly neobsahují jenom matematický popis a vysvětlení různých metrik, ale na reprezentativních datech bakalář vysvětluje jejich použití v praxi. Nezapomíná taky vyzdvihnout účel a slabiny používaných algoritmů.

Součástí zadání diplomové práce bylo vypracování vzorové studie pro účely vyhodnocení kriminálních dat. Bakalář řešení logicky rozdělil na fázi přípravy dat a vytvoření tří různých modelových úloh: nalezení optimálního rozložení hlídek, nalezení souvisejících případů a kapesné krádeže z pohledu lokace a období. V práci podrobně rozbírá způsob přípravy vstupních hodnot, jejich doplnění a nalezení vhodných záznamů pro jednotlivé modelové případy. Na vypracovaném proudu „Příprava dat“ je vidět, jak bakalář vhodně připravil vstupní data, bylo nutné přetypovat špatně rozpoznané atributy a doplnit odvozené data, pro lepší čitelnost následných proudů. Pro případ nalezení optimálního rozložení hlídek bakalář doplnil atributy o časové rozmezí dne, na noční dobu aplikoval algoritmus TwoStep. Modelová úloha nalezení souvisejících případů vloupání do obytných budov dokazuje, že bakalář tématice dolování dat věnoval hodně času a úsilí. Výsledný proud prezentuje, že bakalář nepodcenil přípravu před-kandidátů a na vstupní množinu dat aplikoval dvě různé shlukovací metody: K-means a Kohonenovu mapu, ve třech variantách řazení dat. Úspěšně si poradil taky se sjednocením výsledných shluků.

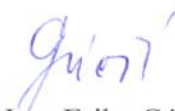
Kladně hodnotím nastudování a pochopení metod dataminingu a následnou implementaci algoritmu K-Means ve vývojovém prostředí .Net. Bakalář vytvořil funkční aplikaci s možností importovat data, zpracovat je pomocí metody K-means a výstup prezentovat v grafu. Implementovanou knihovnu K-means bakalář vhodně rozšířil o další metody.

Písemné zpracování diplomové práce je na vysoké úrovni, doplněné řadou názorných obrázků a rovnic. Vytknout možno jen to, že téma e-learningového kurzu a implementace algoritmu je popisovaná ve více kapitolách bakalářské práce, ideálnější a přehlednější by bylo těmto tématům věnovat souvislou kapitolu.

Diplomovou práci hodnotím známkou

- výborný -

V Liberci dne 6.6.2013


Ing. Erika Gúčíková