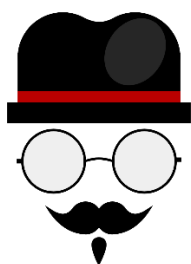




Ročníková projekt

Programování – 3.E



MrCrypto

Autoři: David Pavlův, Filip Dávidík, Bára Siptáková, Ondřej

Záhorský

Škola: Gymnázium, Praha 6, Arabská 14

Kraj: Praha

Konzultant: Mgr. Jan Lána

Praha 2019/2020

Prohlášení

Prohlašuji, že jsem jediným autorem tohoto projektu, všechny citace jsou řádně označené a všechna použitá literatura a další zdroje jsou v práci uvedené.

Tímto dle zákona 121/2000 Sb. (tzv. Autorský zákon) ve znění pozdějších předpisů uděluji bezúplatně škole Gymnázium, Praha 6, Arabská 14 oprávnění k výkonu práva na rozmnožování díla (§ 13) a práva na sdělování díla veřejnosti (§ 18) na dobu časově neomezenou a bez omezení územního rozsahu.

V Praha dne 19. 5. 2020

.....

Anotace

Tato práce se zabývá vývojem ceny Bitcoinu. Aplikace shromažďuje články z populárních internetových novinářských serverů a sociálních sítí zcela automatizovaně. Také poskytuje data z Google Trends a provádí technické analýzy včetně techniky strojového učení, která se snaží porozumět trhu a v průběhu času se učí. Dále pak poskytuje reálnou cenu Bitcoinu a na grafu zobrazuje historické pohyby cen. Výsledkem naší práce je webová stránka, která slouží investorům jako ucelený zdroj informací a server, který data neustále zpracovává, ukládá a obsluhuje klienty.

This work concerns the development prices of Bitcoin. The applications store articles from popular Internet news servers and social networks automatically. More available data from Google Trends and technical analysis, including machine learning techniques that help you understand the market and learn over time. Another offers the real-time price of Bitcoins and booking charts of historical price movements. The result of our work is a website that allows investors as a comprehensive source of information and a server that processes data, stores and serves clients.

Obsah

1	Úvod	5
2	Faktory ovlivňující cenu Bitcoinu	6
2.1	Poptávka a nabídka	6
2.2	Media a veřejnost	6
2.3	Politické změny	6
2.4	Technické vlastnosti	6
3	Vývoj aplikací	7
3.1	Výběr technologií	7
3.2	Backend	7
3.2.1	Historické kurzy	7
3.2.2	Google Trends	8
3.2.3	Indikátory	8
3.2.4	Strojové učení	9
3.3	Frontend	10
3.4	Použité technologie	10
3.5	Výstřižky z webové stránky	10
4	Závěr	12

1 Úvod

V současné moderní době vznikají zajímavé a průlomové technologie každou chvíli. Některé se vypracují a stanou se nezbytnou součástí každodenního života a některé se ukáží jako nepotřebné či nedůležité. Bitcoin zcela určitě patří, i když tomu tak nebylo do nedávna, do té šťastnější skupiny; důkazem je tomu neustálý mediální kontext a jeho rostoucí využití.

Jedná se o nejpopulárnější decentralizovanou kryptoměnu. Roku 2008 ji vytvořil člověk nebo skupina lidí podepsaná jako Satoshi Nakamoto (1). Jeho cena se od roku 2010 ke dnešnímu dni (1. 3. 2020) zvýšila téměř o 10 milionů procent. Avšak i přesto, že je Bitcoin tak populární, v porovnání s ostatními aktivy je stále extrémně volatilní ¹ a představuje velmi riskantní investici. Jsou tací, kteří dosáhli neuvěřitelných zisků a svým příběhem motivují ostatní, ale jsou také tací, kteří na tom nemile prodělaly. Bitcoin se definuje jako plně funkční kryptoměna, ale těžko ho můžeme srovnávat s klasickými měnami, kterou jsou regulovány, běžně používané a velice stabilní. Tím pádem mnozí chápou volatilitu Bitcoinu jako příležitost, jak vydělat peníze a zkoušejí své štěstí. Jejich cíl je koupit a prodat draž; takové krátkodobé spekulace, i když si to jejich příznivci někdy neuvědomují, paradoxně přispívají volatilitě.

2 Faktory ovlivňující cenu Bitcoinu

Hlavním krokem k jeho stabilizaci je běžnost používání, pokud se s ním například bude platit i v pekárnách, supermarketech apod. stejně jako s eury, tak faktory, které ovlivňují cenu nakonec neovlivní cenu tolik protože nebudou hrát vysokou roli v porovnání s transakcemi lidí.

2.1 Poptávka a nabídka

Objem Bitcoinů, které lze vytěžit, je předem daný, takže jde o podobnou komoditu jako zlato nebo ropa (2). Pokud je větší poptávka, cena roste. Pokud nabídka roste, cena klesá. V současné době je v koloběhu 18 246 012.5 Bitcoinu. A maximum, které lze vytěžit je 21 milionu, takže zbývá 2 753 987.5 Bitcoinů. Na paměti je dobré mít, že mnoho z již vytěžených Bitcoinu jsou nadobro ztraceny (odhaduje se 3-4 miliónu Bitcoinů), protože jejich majitelé ztratili přístup ke svým elektronickým peněženkám. (3)

2.2 Media a veřejnost

Tento faktor silně souvisí s poptávkou a nabídkou. Názory vlivných lidí na sociálních sítích a media obecně ovlivňují masy lidí. A masy lidí ovlivňují cenu Bitcoinu. Jedná se možná o ten nejdůležitější a nejlépe sledovaný faktor. Často se stává, že se do Bitcoinu rozhodla investovat nějaká firma, že jej začíná akceptovat jako platidlo nějaký velký obchod a podobně. (2)

2.3 Politické změny

Politické události a státní regulace můžou výrazně ovlivňovat kurz Bitcoinu, a to oběma směry. Když se Čína, největší trh s kryptoměnami, rozhodla omezit využívání Bitcoinu a uzavřela několik jeho burz, jeho cena klesla dolů. Když naopak japonská vláda uznala Bitcoin jako oficiální platidlo, cena několik měsíců dramaticky rostla. (2)

2.4 Technické vlastnosti

Mezi technické faktory patří například tzv. halving. Jedná s fází, kterým si prochází Bitcoin každý 210 000. vytěžený blok způsobí, že se počet nových dodávaných Bitcoinů sníží o 50%, a tím se také sníží i odměna pro těžaře o 50%. To může mít za následek, že se těžařům nevyplatí těžit Bitcoin, protože cena elektřiny pro výpočet stojí více než by odměna v Bitcoinech. (4)

Vliv na cenu Bitcoin má také vznik nové kryptoměny pomocí forku Bitcoinu. Takovým kryptoměnám se obecně přezdívá litecoiny a jsou konkurencí pro Bitcoin samotný. (5)

3 Vývoj aplikací

Cílem našeho projektu bylo vytvořit ucelený nástroj pro investory, který by jim pomohl usoudit, zda je dobrý čas na investování do Bitcoinů.

Rozhodli jsme naprogramovat server, který bude shromažďovat, zpracovávat a ukládat data. A pak je posílat na požadavek. A dále jsme naprogramovali i webovou stránku, která bude data stahovat ze serveru a zobrazovat do přehledných grafů a listů.

3.1 Výběr technologií

Většina práci jsme dělali v Javascriptu. Jedná se o objektově orientovaný programovací jazyk, který nabízí vše k vývoji interaktivních webových stránek.

Nabízí 3 frameworky k vytváření UI: Vue.js, Angular.js a React.js. My jsme si vybrali **React.js**, protože je nejlépe aktualizován.

U serveru jsem se rozhodli také pro Javascript. Pomoc **Node.js** (napsané v C, C++, JavaScript) (6), což je prostředí umožňující spouštět JavaScript kód mimo webový prohlížeč. To nám umožnilo psát frontend i backend v Javascriptu. Pro jednodušší obsluhování a fungování serveru jsme použili framework **Express.js** a pro ukládání NoSQL databázi **MongoDB**.

Tento výběr technologií je často používaný a kvůli jednotnému programovacím jazyku velice pohodlný. Obvykle se nazývá zkratkou MERN Stack (první písmeno z každé technologie).

3.2 Backend

Backend má krom stahování a obsluhování klientů zpracovávací funkci. Aby tyto data nemusel zpracovávat každý klient zvlášť, stačí aby tyto data zpracoval jednou právě server a výsledek posílal. Tímto šetříme výkon.

Efektivní optimalizací posílání dat se stalo jejich předchozí stringování. Data mají totiž v databázi formát .json. Před každým požadavkem data vždy zkomprimujeme (dochází k ušetření místa v řádech desítek %) z .json do prostého textu. Až data dorazí, klient si sám zase naparsuje do pohodlného .json. Tímto postupem šetříme požadavky na rychlost sítě.

3.2.1 Historické kurzy

Mít přístup k aktuálním a historickým kurzům zcela zdarma se ukázalo jako komplikovaný cíl.

Mnohé API² zdarma sice poskytují data, ale omezují je na nízkou aktualizaci dat a vysokou odezvu. Dále bylo potřeba zjistit kvalitu dat z jednotlivých API, některá totiž obsahovala časová okna úplně bez dat.

Po otestování těch nejpopulárnějších se ukázalo, že každá API má nějaké nedostatky a nejlepší bude použití nějaké kombinace. První z nich je Coinbase Pro. Poskytuje rychle se aktualizující

² API: rozhraní pro programování aplikací.

data v různých intervalech a částech. Avšak v některých měsících data vykazovaly ztráty. Tuto nevýhodu jsme vyřešili s kombinací BitcoinAverage. Jejich data se sice aktualizují každé asi 3 dny, alespoň v jejich datech nejsou žádné ztráty.

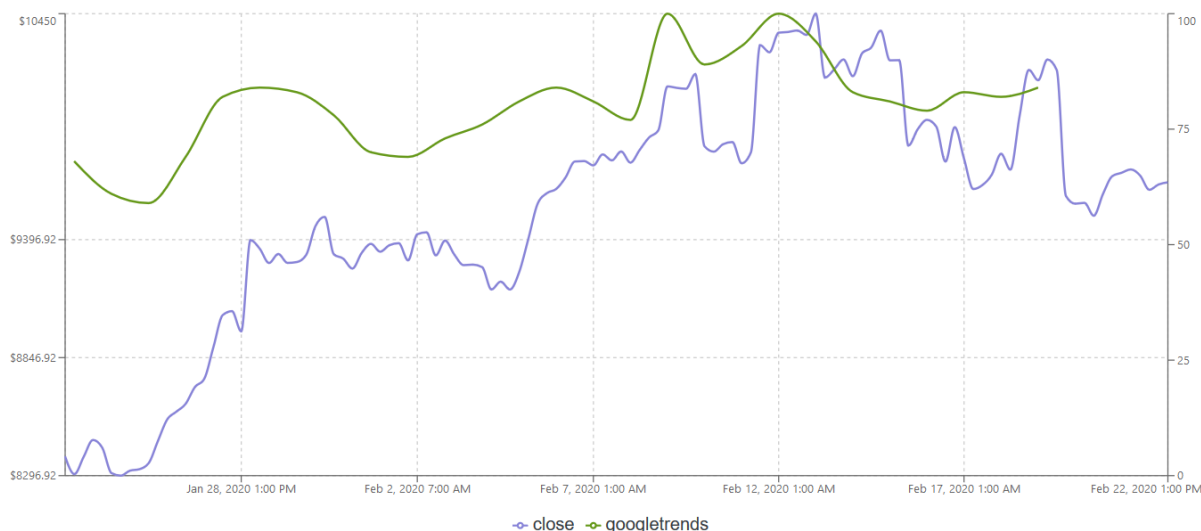
U obou bylo potřeba ošetřit omezení požadavků. U Coinbase Pro je to maximální počet kurzů na 300. Například denní kurzy za celý rok jsme museli rozdělit na 2 požadavky po 300 a 65. To se provádí na pozadí zcela automatizovaně díky aplikované a chytré funkci.

Naše API je teda kombinací dvou zdrojů, které si kompenzují své nedostatky.

3.2.2 Google Trends

API jsme také vylepšili o zajímavou funkci od Googlu, která poskytuje zájem lidí o klíčové slovo (v našem případě „bitcoin“). Data neobsahují přesná čísla, ale poměry v daném úseku, což pro zobrazení v grafu není žádný problém; postačí jako ukazatel zájmu.

Data se stejně jako kurzy aktualizuje průběžně.



Obrázek 1.- korelace ceny a zájmu

Na obrázku 1. je výrazně vidět jak cena Bitcoinu v průběhu 1 měsíce koreluje se zájmem lidí. Tento zajímavý ukazatel určitě stojí za zaobírání jak pro investora, tak i pro analytiku v oboru. Za zmínku stojí také říct, že zájem v Google Trends, lze výkonnými počítači lehce zmanipulovat.

3.2.3 Indikátory

Jedná se o ukazatele, které se chovají jako trend. Pokud jde cena nižší než tento indikátor, znamená to, že je cena pod dlouhodobým trendem. Obecně pokud je cena blíže k indikátoru, dá se po překřížení očekávat růst nebo pokles; v závislosti na směru křížení.

Jednoduchý klouzavý průměr je specifický tím, že přiřazuje všem hodnotám stejnou důležitost, což se dá vnímat jako nevýhoda. Pokud je poslední hodnota extrémní, může se po přidání další

hodnoty průměr rapidně změnit opačným směrem, což nemusí znamenat, že nastal prudký vzrůst či spád v ceně. Počítáme jej tedy $(cena\ n + cena\ n-1 + cena\ n-2)/n$.

U váženého klouzavého průměru přiřazujeme každé hodnotě jinou váhu sestupně od hodnoty bližší momentální hodnotě.

Exponenciální klouzavý průměr dává vyšší váhu posledním hodnotám. Počítáme jej tedy: $(EMA\ n-1) + [x * (cena\ n - EMA\ n-1)]$, kdy $x = 2/(n+1)$.

Triangulární klouzavý průměr (TMA) je jednoduchý klouzavý průměr jednoduchého klouzavého průměru, výsledek TMA bývá vyhlazená křivka původního jednoduchého klouzavého průměru.

Bollingerova pásma používají klouzavé průměry tak, že vykreslují čáry v určité vzdálenost nad a pod klouzavý průměr. Bollingerova pásma tvoří tři křivky. Středová zelená křivka zobrazuje klouzavý průměr, Kolem středu je vytvořena obálka proměnlivé šířky (modrá). Obálku tvoří násobek r směrodatných odchylek n posledních zavíracích kurzů.

3.2.4 Strojové učení

Strojové učení bylo do nedávna záležitostí především v jazyce python. Díky velkému množství peněz investovaných do vývoje tohoto oboru, ale došlo k rozšíření v mnoha směrech (7).

Jedním z nich je TensorFlow.js, knihovna disponující neurální sítí. Neurální síť potřebuje co největší množství dat k učení. V našem případě s databází to není však problém.

Formát vstupu do sítě jsme museli rekonstruovat do formátu, který je vhodný a nejvíce efektivní. Normalizování dat pomocí škálování se ukázalo jako špatný nápad, protože ve vybraném časovém úseku se cena často pohybovala v rozmezí maximálně jedné stovky dolarů, což u aktuálního kurzu (ke dni 1. 3. 2020 téměř \$8000) nemá dostatečnou páku na penalizaci (learning rate) neurální sítě. Nejnižší hodnoty $loss^3$ při 400 epochs⁴ s přibližně 500 vstupy jsme dosáhli přeformátováním dat na sekvence 0s a 1s. Nastaveno na 0 pokud ve srovnání s předchozím dnem cena klesla, a 1 pokud vzrostla. Pro výstupy jsme nastavili hodnotu následující dne. Pro velikost okna 5 (nakonec jsme zvolili velikost 20) by vstup mohl vypadat například takto: [0,0,1,0,1] a pro výsledek [1]. V „překladu“ to znamená, že první dva dny cena klesala, 3. vzrostla, 4. klesla, 5. vzrostla a 6. den, kterým je výsledek, cena vzrostla. Neurální síť tyto vstupy projede a pokusí se zjistit nějakou spojitost a pravidelnost mezi vstupy a výsledky. Nezbytnou součástí je míchání dat a vybalancování dat, pokud totiž budou mít trénovací data převážně výsledky vzrůstu, tak i predikce vytrénované sítě budou spíše sklouzávat k růstu. Dále bylo potřeba data otestovat, na reálných datech; na 3 pokusech vždy o velikosti 50 testovacích dat jsme dosáhli úspěšnosti 55%, 46%, 56% (v průměru 52,33%), což jsou velice uspokojivé výsledky.

³ Loss je hodnota na určení přesnosti vytrénované neurální sítě. Čím nižší, tím přesnější.

⁴ Epoch je počet, kolikrát jedny data použijeme pro vytrénování neurální sítě. Pokud máme 10 vstupů a epochs nastavené na 30, tak neurální síť projde 300 dat.

3.3 Frontend

Je něco, čemu jsme kladli veliký důraz; stránka vypadá profesionálně a nabízí příjemný vzhled, logickou navigaci strukturu, interaktivitu s uživatelem a další.

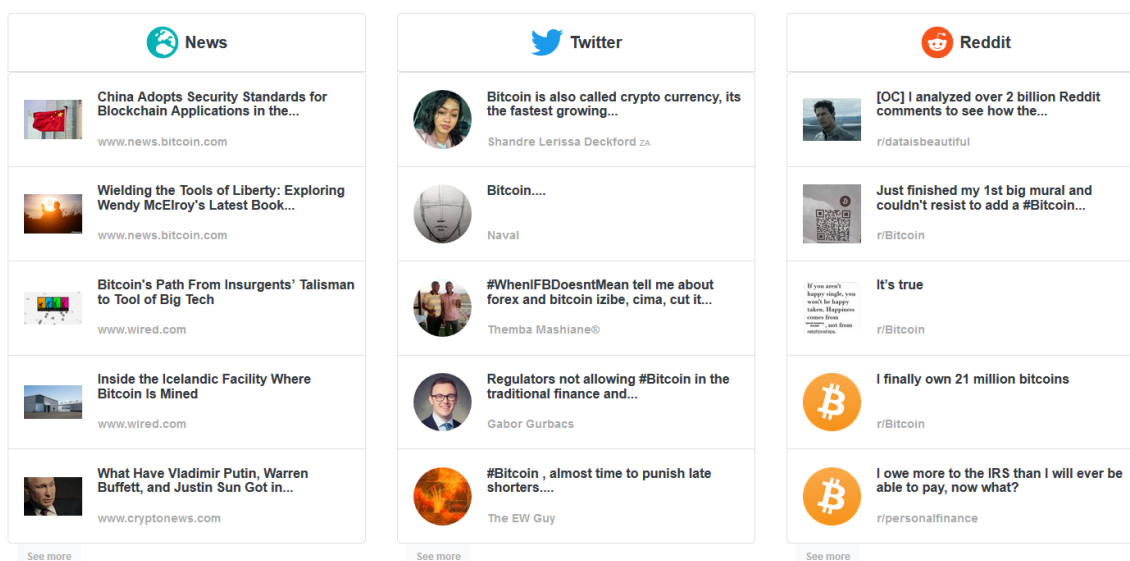
3.4 Použité technologie

Stránka je naprogramovaná pomocí javascriptového frameworku React.js. Dále jsme použili knihovnu Redux, který se stará o uložená data (states) a aktualizuje jen ty komponenty, který to opravdu potřebují. Například cena Bitcoinu se aktualizuje každou minutu, avšak to neznamena, že se musí aktualizovat něco jiného; takové vypracování ve webové stránce, která neustále něco aktualizuje serveru, nesmělo rozhodně chybět.

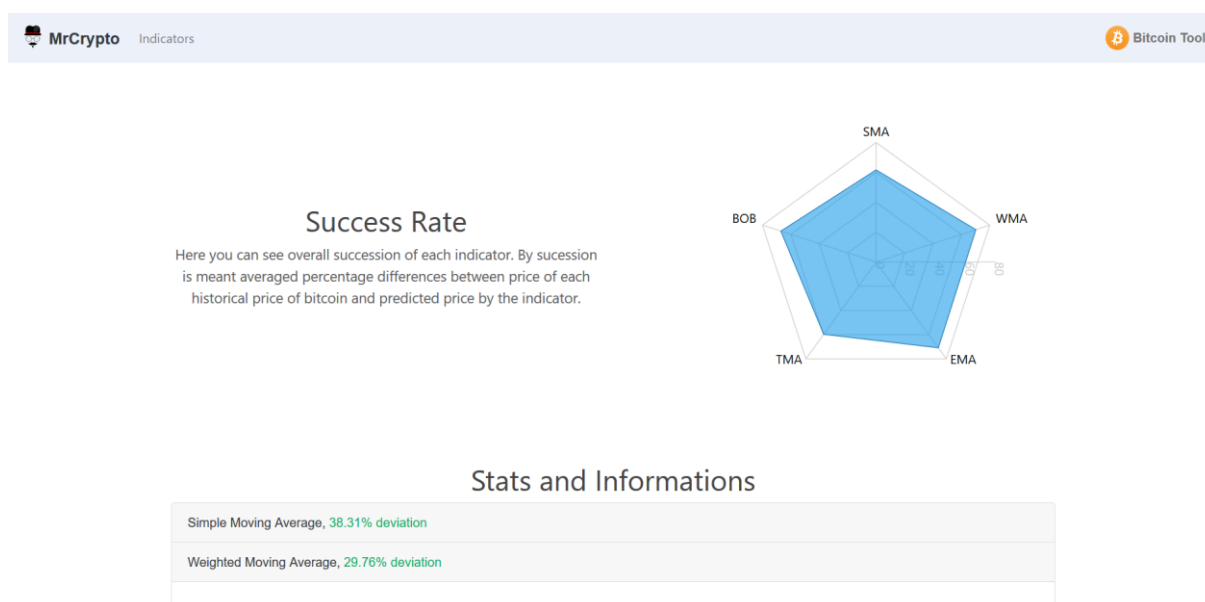
3.5 Výstřižky z webové stránky



Obrázek 2. - úvodní pohled na stránku



Obrázek 3. - struktura zobrazení článků o Bitcoinu



Obrázek 4. - přehled o indikátorech a jejich úspěšnosti

3.6 Jak používat webovou stránku

Na úvodní straně webové stránky najdete aktuální novinky ze světa, ohledně Bitcoinu. Na každý článek lze kliknout, což vás přeměruje na zdroj článku. Také zde vidíme graf, který má interaktivní tlačítka: tlačítka na levo jsou pro spuštění indikátorů a GoogleTrends, na pravo jsou pro určení časové periody. Graf v základu zobrazuje historická data hodnot kryptoměny Bitcoin. Na druhé straně “indicators“ lze vidět úspěšnost indikátorů a když si rozkliknete tlačítka pod “stats and informations“, tak najdete i vysvětlení konkrétního indikátoru. Nakonec zde najdete předpověď naší neurální sítě.

Závěr

Tato ročníková práce má mnoho zajímavých podkategorií, které skrývají vysoký potenciál. A i když jsme zadání práce splnili, tak je tu prostor pro další vylepšení a kroky k lepšímu pochopení, jak kryptoměny fungují a reagují různé podněty.

Budoucnost aplikace vidíme spíše ve vzdělávacím směru. Určitě může být aplikace vnímaná jako reálný simulátor a být obohacena o kupování a prodávání. A tím sloužit začínajícím investorům jako pískoviště pro jejich budoucí reálné investice.

4 Bibliografie

1. [Online] <https://cs.wikipedia.org/wiki/Bitcoin>.
2. [Online] <https://www.investicniweb.cz/pet-faktoru-ovlivnujicich-cenu-bitcoinu/>.
3. [Online] <https://www.buybitcoinworldwide.com/how-many-bitcoins-are-there/>.
4. [Online] <https://blockgeeks.com/guides/bitcoin-halving/>.
5. [Online] <https://connect.zive.cz/clanky/kryptomenove-forky-jak-a-proc-dochazi-k-rozdeleni-kryptomen/sc-320-a-192682/default.aspx>.
6. [Online] <https://en.wikipedia.org/wiki/Node.js>.
7. [Online] <https://www.techworld.com/picture-gallery/data/tech-giants-investing-in-artificial-intelligence-3629737/>.

Seznam obrázků

Obrázek 1.- korelace ceny a zájmu	8
Obrázek 2. - úvodní pohled na stránku	10
Obrázek 3. - struktura zobrazení článků o Bitcoinu	11
Obrázek 4. - přehled o indikátorech a jejich úspěšnosti.....	11

