

Mendelova univerzita v Brně
Provozně ekonomická fakulta

Webová aplikace pro krátkodobou prognózu Cash Flow

Diplomová práce

Vedoucí práce:

Ing. Michael Štencel, Ph.D.

Řešitel:

Bc. Marek Vacek

Brno 2012

Zde vložit zadání práce!

Chtěl bych tímto poděkovat rodině za podporu při studiu a vedoucímu své práce Ing. Michaelu Štenclovi, Ph.D. který mi aktivně poskytoval tolik potřebné rady a připomínky k této práci.

Prohlašuji, že jsem tuto práci vytvořil sám, s použitím zdrojů uvedených v seznamu literatury.

V Brně, dne 3. ledna 2012

.....

Abstract

Vacek, Marek. *Application for short-term cash flow forecasting*. Diploma thesis. Brno, 2012.

This thesis focuses on design of an application for automatic personal cash flow forecasting. The paper examines the current state and proposes statistical methods, by which the prediction is feasible. There are also included results of tests on data acquired in a survey. The end of the thesis discusses possibility of usage of neural networks and ways to implement and finance the application.

Abstrakt

Vacek, Marek. *Webová aplikace pro krátkodobou prognózu Cash Flow*. Diplomová práce. Brno, 2012.

Tato práce se zabývá návrhem aplikace, jež bude provádět automatickou prognózu osobních finančních toků. Práce zkoumá současný stav a navrhuje statistické metody, na jejichž základě lze provádět automatizované prognózování osobních finančních toků. Jsou zde uvedeny výsledky testování metody nad daty získanými v anketě. V závěru pak práce diskutuje použití metod umělé inteligence, možnosti vytvoření, nasazení a financování této aplikace.

Obsah

1	Úvod	8
2	Cíl práce	9
3	Analýza současného stavu	10
3.1	Způsoby sestavení a udržování finančního plánu	10
3.2	Dostupné aplikace a software pro plánování cash flow	10
3.3	Možnosti a limitace predikce finančních toků	11
4	Metody práce a metodika	12
4.1	Základy osobního finančního plánování	12
4.2	Potřeba automatizované předpovědi cash flow	12
4.3	Ekonomické časové řady a jejich vlastnosti	13
4.3.1	Časové řady a jejich dekompozice	15
4.3.2	Sezónní a cyklické prognózy ekonomických řad	16
4.3.3	Trend	16
4.3.4	Sezónnost	18
4.3.5	Korelace Dat	19
4.3.6	Bodový diagram	20
4.3.7	Nelinearita	20
4.4	Cash flow	21
4.4.1	Použitá struktura cash flow	22
4.4.2	Teoretický základ vykazování cash flow	22
4.4.3	Metody vykazování cash flow	23
4.4.4	Čistá přímá metoda	23
4.4.5	Nepřímá metoda	23
4.5	Forma predikce	24
4.6	Měření přesnosti predikce	24
4.7	Výdajové a příjmové typy	24
4.8	Metodika návrhu aplikace	25
5	Automatizace predikce cash flow	27
5.1	Rozložení cash flow na jednotlivé kategorie	27
5.2	Data pro testování algoritmu a složení vzorku	27
5.2.1	Respondenti	28
5.2.2	Časové řady	28

5.3	Datové případy, jejich definice a extrapolace	29
5.3.1	Datový případ – silná sezónnost	29
5.3.2	Datový případ – stejná data a schodky	35
5.3.3	Datový případ – nepoužívaná řada	36
5.3.4	Datový případ – velice krátká řada	37
5.3.5	Datový případ – řada bez historických dat	38
5.3.6	Datový případ – různé hodnoty bez periodicity	38
5.4	Korelace dat	40
6	Vlastní aplikace	42
6.1	Specifikace požadavků	42
6.1.1	Účel a využití aplikace	42
6.1.2	Základní popis aplikace	42
6.1.3	User stories	42
6.1.4	Uživatelé	43
6.1.5	Vkládání dat k budoucím pohybům	43
6.1.6	Prezentace a vysvětlení výsledků uživateli	44
6.1.7	Uživatelské rozhraní	45
6.1.8	Bezpečnostní požadavky	45
6.1.9	Architektura a technické řešení aplikace	45
6.2	Databázový návrh	46
6.3	Podoba aplikace	46
6.4	Implementace algoritmu predikce	47
6.5	Test úspěšnosti predikce	48
6.6	Porovnání predikcí	48
6.7	Finanční aspekty provozu aplikace	49
7	Diskuse	52
8	Závěr	53
9	Použitá literatura	54

1 Úvod

Plánování osobních financí je jeden z nejvíce zanedbávaných oborů finančnictví vůbec. Ve firemním i veřejném sektoru se na sběr dat, analýzu a plánování finanční strategie vynakládají obrovské prostředky i úsilí. Oproti tomu osobní finance v tomto ohledu často nedostávají takto profesionální péči. Každý svůj rozpočet nějakým způsobem plánuje, dělá to ovšem víceméně intuitivně, aniž by měl nějaké hlubší znalosti z této problematiky. Takové plány navíc většina lidí vytvoří pouze v situaci potřeby větších výdajů, například při nákupu nemovitosti či vozidla, ovšem tehdy už jen zjistí, že nejsou na takovou situaci připraveni. Někteří lidé ve zlomových okamžicích života vyhledají pomoc finančního poradce, tito jsou ale málokdy schopni opravdu zevrubně analyzovat situaci klienta a poskytnout mu potřebná doporučení. V nemálo případech jsou takovíto poradci pouze prodejci finančních produktů a situaci klienta zmapují pouze do míry, která je potřebná pro jejich účel.

Pro efektivní využití peněz po celý život je proto třeba mít permanentně sestaven celoživotní finanční plán a neustále jeho podobu měnit s tím, jak se mění osobní požadavky a životní situace. Toto ovšem není jednoduchou záležitostí. Uložení financí na více místech, zamčení v nejrůznějších produktech, půjčky, splátky, pohledávky, proměnlivá hodnota peněz v čase, rostoucí náklady v čase, proměnlivé příjmy a mnoho dalších proměnných a nástrah, toto vše dělá problematiku plánování osobních financí nesmírně složitou. Pokud má někdo naspořen dostatek finančních prostředků, pokryje nastalou situaci ze svých finančních rezerv. Takovéto osoby se nedostávají do existenčních potíží, ale často jim uniká zisk z nevyužitých prostředků. Odhadnout správnou výši rezervních zásob je totiž pro výše uvedené důvody velice těžké. Pak jsou zde sociální vrstvy, kde jsou finance spotřebovávány tzv. za chodu a pro ty jsou jakékoliv nečekané výdaje potenciálně existenčním problémem. V případě těchto sociálních vrstev je dobré finanční plánování nutností a v ideálním případě vede k vytvoření vyváženého rozpočtu, který jim z této svízelné situace pomůže. Proto doufám, že výstup této práce pomůže všem lidem, kteří nemají dokonalý přehled o svých financích.

2 Cíl práce

Cílem práce je navrhnout aplikaci pro krátkodobou prognózu cash flow pro použití při plánování osobních financí. Práce nejprve prozkoumá možnosti statistické predikce osobních peněžních toků v krátkém období za použití automatizovaného systému. Na základě těchto poznatků bude navržena podoba aplikace využívající zvolené metody. Práce otestuje navržené metody predikce nad testovacími daty získanými v anketě. V závěru pak práce diskutuje použití metod umělé inteligence, možnosti vytvoření, nasazení a financování této aplikace.

3 Analýza současného stavu

V současné době při potřebě osobního finančního plánu využívají různé nástroje elektronické i neelektronické. V této sekci proto diskutujeme dostupné nástroje.

3.1 Způsoby sestavení a udržování finančního plánu

Způsoby sestavení a udržování finančního plánu můžeme rozdělit například podle nástroje, který je v procesu využíván.

- Tzv. „Tužka a papír“ – jedná se o použití neelektronických prostředků pro celý proces plánování.
- Tabulkové procesory – v dnešní době jeden z nejdostupnějších a zároveň nejpoužívanějších elektronických nástrojů k finančnímu plánování. Jejich velká dostupnost a všeobecná znalost vede k jejich oblíbenosti. Jedná se o skvělý nástroj, který jednoduše provádí výpočty v reálném čase a nabízí i velkou variabilitu. Díky jeho všeobecnosti ovšem chybí pokročilejší nástroje k podpoře plánování. Navíc je uživateli nenavádí správnou cestou a ten musí proto znát správný postup k sestavení plánu.
- Software a aplikace pro finanční plánování – jedná se o použití o dedikovaného software pro hospodaření s penězi. Ten uživatele postupně navádí k sestavení finančního deníku. V této práci se budeme zabývat především tímto způsobem zpracování finančního plánu.

3.2 Dostupné aplikace a software pro plánování cash flow

V dnešní době je dostupných mnoho aplikací umožňující nějakou formou sledovat a zpřehlednit osobní finance. Některé aplikace umožňují nějakou formou plánovat budoucí cash flow, ale je to vždy jen okrajová funkce celého balíku a málokdy umožňuje opravdu efektivně plánovat svoje peněžní toky. Žádná z nalezených aplikací ovšem automaticky nemodeluje budoucí vývoj dostupných peněžních prostředků. Nejrozšířenější aplikace jako Fortora Fresh Finance, Moneydance, Quicken, GnuCash, Mint.com a Thrive se zaměřují především na analýzu stávající situace a do budoucna umožňují jen stanovení jakéhosi fixního měsíčního rozpočtu, jehož dodržování poté kontrolují.

3.3 Možnosti a limitace predikce finančních toků

Z povahy chování spotřebitelů je zřejmé, že ne všechny položky, ať už příjmové či výdajové, se nedají spolehlivě předpovědět. Vysoce nepravidelné výdaje a příjmy, stejně jako nové položky, pro které neexistuje historie v datech žádný model nemůže předpovědět. Proto jakákoliv aplikace, která bude modelovat budoucí vývoj cash flow, musí umožnit uživateli doplňovat data pro budoucí období. Stejně tak musí nabízet možnost u některých druhů příjmů a výdajů modelování neprovádět (například pokud se uživatel rozhodne již za danou věc peníze nevydávat). U vysoce nepravidelných výdajů model nanejvýše určí průměrnou měsíční částku.

Dále je nemožné předvídat s užitečnou přesností finanční toky na příliš období dopředu. Makroekonomické síly stejně jako měnící se životní situace osoby, jejíž cash flow modelujeme, způsobují, že s narůstající vzdáleností do budoucna, ztrácí předpověď přesnost. Pro praktické použití proto většina odborných poradců sestavuje plány cash flow na delší období než jeden rok. Proto i v této práci se zaměříme především na krátkodobé modelování do období jednoho roku.

4 Metody práce a metodika

4.1 Základy osobního finančního plánování

Finanční plánování je proces, skrz který můžeme navrhnout plán pro splnění všech očekávaných i neočekávaných potřeb v životě. Záměrem je, aby byla přijata nezbytná opatření k zajištění všech potřeb k splnění všech cílů a aby byl jedinec připraven také na většinu nepředvídaných událostí. [2]

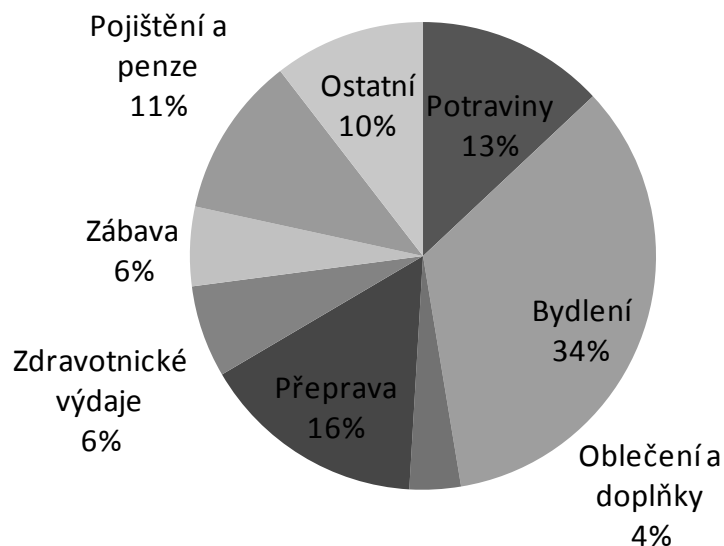
Fáze plánování dělíme do 5-ti základních kroků.

1. **Zhodnocení osobní finanční situace.** A s tím související sběr dat. Může být posouzena například zjednodušenou verzí rozvahy a výkazem zisků a ztrát. Osobní rozvahou rozumíme sepsání hodnoty majetku spolu s osobními závazky. Výkaz zisku a ztrát vytvoříme uvedením příjmů a výdajů.
2. **Stanovení cílů.** Je důležité stanovit například věk odchodu do důchodu, nákupu nemovitosti apod.
3. **Vytvoření plánu.** Finanční plán podrobně popisuje jak dosáhnout předem stanovených cílů. To zahrnuje doporučení jako snížení nákladů, zvýšení příjmů nebo investování peněz.
4. **Uskutečnění.** Uskutečnění plánu vyžaduje disciplínu a vytrvalost. Mnoho plánů nefunguje právě proto, že nejsou dodržovány.
5. **Monitorování a přehodnocování.** S průběhem času je třeba osobní plán upravovat a posuzovat vzhledem ke skutečnosti a změnám v prioritách jedince.

4.2 Potřeba automatizované předpovědi cash flow

Při sestavování plánu cash flow zvládne i nezkušený jedinec jednoduše stanovit do budoucna objem pravidelných položek neměnných, ovšem hodnotu variabilních položek dokáže určit, jen pokud bude pečlivě sledovat svoje výdaje a sestaví si jejich strukturu, dostane se pak k průměrné částce, kterou měsíčně vydává. Ta ovšem v průběhu času může vykazovat silné sezónní výkyvy. Při sestavování cash flow pak použije přinejlepším pouze průměrnou hodnotu výdajů a nezohlední tak sezónnost výdajů ani korelaci s jinými položkami.

Struktura výdajů běžné domácnosti



Obr. 1: Struktura výdajů běžné domácnosti v USA [3]

Z výzkumu osobních výdajů amerického úřadu práce je zřejmý vysoký podíl variabilních položek, a to obzvláště ve výdajové části osobního rozpočtu.[3] Proto plán sestavený bez zohlednění všech těchto vlivů může mezi jednotlivými měsíci vykazovat značné nepřesnosti.

Analýza sezónnosti a korelace se všemi vstupy je ovšem velice náročná a musí se udělat při každé změně rozpočtu. Je proto vhodné aby takovýto problém byl řešen v reálném čase elektronickou aplikací. Aplikace by proto měla jedinci vymodelovat předpokládaný vývoj jeho cash flow v budoucnu na základě dat z minulých období při zohlednění všech vlivů.

4.3 Ekonomické časové řady a jejich vlastnosti

Důležitým úkolem statistických analýz ekonomických jevů je zkoumání jejich dynamiky. Empirická pozorování v ekonomické oblasti jsou často uspořádána do časové řady. Ekonomickou časovou řadou se rozumí řada hodnot jistého věcně a prostorově vymezeného ekonomického ukazatele, která je uspořádána v čase směrem od minulosti do přítomnosti. [1]

Ekonomické časové řady lze klasifikovat podle typu ukazatele, který se dělí, na intervalové a okamžikové. Intervalové časové řady jsou řadami ukazatelů, jejichž hodnoty závisí na délce časového intervalu sledování. Typickými intervalovými ukazateli jsou extenzitní ukazatele, jejichž příkladem může být objem výroby, spotřeba surovin atd. Okamžikové časové

řady jsou řadami ukazatelů, jejichž hodnoty se vztahují k jistým časovým okamžikům. Hodnoty takových ukazatelů nezávisí na délce časového intervalu sledování. Příkladem okamžikového ukazatele je počet neumístěných uchazečů o zaměstnání evidovaných na úřadech práce k určitému datu. [5]

Klasifikaci ekonomických časových řad lze provést také podle délky intervalu sledování hodnot. Dlouhodobé časové řady mají hodnoty sledované v ročních či delších časových úsecích, hodnoty krátkodobých časových řad se sledují v úsecích kratších než jeden rok, a vysokofrekvenční časové řady mají hodnoty sledované v úsecích kratších, než je jeden týden. [5]

Lze pozorovat, že zejména s druhou klasifikací souvisí tvar ekonomických časových řad, např. čím je interval sledování delší, tím jsou časové řady vyhlazenější. Tato skutečnost však vyplývá z typického rysu časových řad – časové „svázanosti“ jejich klíčových hodnot. Na rozdíl od průřezových dat, má u časových řad pořadí hodnot klíčový význam. Způsob jakým na sebe jednotlivé hodnoty v časových řadách navazují, určuje jejich tvar a charakteristické vlastnosti. [5]

Ekonomické časové řady jsou charakteristické: a) trendem, b) sezónností, c) podmíněnou heteroskedasticitou, d) nelinearitou a e) společnými vlastnostmi více časových řad, např. tzv. společným trendem. Tyto vlastnosti se u časových řad neobjevují zpravidla najednou. Jejich přítomnost závisí na typu časové řady, např. sezónnost se objevuje u krátkodobých časových řad, podmíněná heteroskedasticita u vysokofrekvenčních časových řad. [1]

V empirických analýzách se někdy ekonomické časové řady logaritmičsky transformují. Důvodů pro transformaci je několik. Některé ekonomické časové řady jsou charakteristické exponenciálně se vyvíjejícím trendem a logaritmičská transformace znamená jeho linearizaci. Touto transformací se současně časová řada stabilizuje z hlediska variability. V případě finančních časových řad, tj. časových řad cen a funkcí, se vychází z předpokladu, že cena nemůže být záporné číslo, předpokládá se tedy, že by hodnoty těchto časových řad mohly být generovány logaritmičko-normálním rozdělením. Jak je známo, logaritmus náhodné veličiny s logaritmičko-normálním rozdělením má rozdělení normální. Při konstrukci ekonometrických modelů se často vychází z teoretických ekonomických modelů, které jsou v exponenciálním tvaru. Jejich linearizace se dosáhne logaritmováním, do lineárního modelu tedy musí vstoupit logaritmičsky transformované časové řady. [1]

4.3.1 Časové řady a jejich dekompozice

Časovou řadou budeme rozumět posloupnost věcně a prostorově srovnatelných pozorování, která jsou jednoznačně uspořádaná z hlediska času, a to zpravidla chronologicky. [5] Pomocí analýzy časových řad chceme především určit model, který popisuje, jaký vliv má časový faktor na utváření posloupnosti sledovaného jevu, a dále také pak využít určeného modelu k předpovědi budoucího vývoje systému. [4]

Dekompozicí časové řady rozumíme rozložení časové řady na složky, a to na složku trendovou, sezónní, cyklickou a reziduální (zbytkovou, iregulární). [1]

Trendem rozumíme dlouhodobou tendenci v chování pozorovaného ukazatele. V průběhu sledovaného období můžeme sledovat dlouhodobý růst, dlouhodobý pokles nebo mohou hodnoty kolísat kolem určité hodnoty, pak hovoříme o časové řadě s konstantním trendem. [1]

Sezónní složka zobrazuje pravidelně se opakující změny v časové řadě, které probíhají v jednom kalendářním roce a každý rok se opakují. Příčinami sezónních změn jsou nejčastěji vlivy změny ročních období nebo také různé společenské zvyklosti (výplata mezd v určitou dobu, svátky, dovolené, apod.). [1]

Cyklická složka představuje kolísání okolo trendu s periodou delší než jeden rok. Nikdy cyklická složka nebývá považována za samostatnou složku, ale bývá součástí trendu. Délka cyklu i intenzita jednotlivých fází bývá často proměnlivá. Cyklická složka vzniká z dlouhodobých vnějších vlivů např. změny klimatu. Typickým představitelem této složky je tzv. obchodní cyklus, jehož délka se pohybuje v rozmezí 5 - 7 let. [1]

Reziduální složka zbývá v časové řadě po eliminaci složky trendové, sezónní a cyklické. Je tvořena náhodnými pohyby v průběhu časové řady, obvykle také pokrývá chyby vzniklé při měření údajů a při samotné analýze řady (např. zaokrouhlování). [5]

Pro vlastní tvar rozkladu uvažujeme obvykle tyto dva základní typy:

- aditivní dekompozice, kde jsou jednotlivé složky uvažovány ve skutečných absolutních hodnotách a jsou měřeny v jednotkách řady,
- multiplikativní dekompozice, kde je většinou pouze trendová složka uvažována ve své absolutní hodnotě a ostatní složky jsou uvažovány v relativních hodnotách vůči trendu.

4.3.2 Sezónní a cyklické prognózy ekonomických řad

K prognózám ekonomických řad existují dva hlavní přístupy. Odhad budoucích dat může být založen na analýze faktorů, které mají na danou řadu vliv. V takovémto případě mluvíme o vysvětlující metodě. Jiným přístupem je pak extrapoláční metoda, která analyzuje chování dat v čase. Například důvěra v to, že zvýšení poptávky po oblečení pro panenky je důsledkem nedávné reklamní kampaně a ne tím že se právě blíží období vánoc, dobře ilustruje rozdíl mezi těmito dvěma filozofiemi. Je možné, že oba tyto přístupy vedou k vytvoření přesné a užitečné prognózy ekonomické řady, ovšem metoda vysvětlující se velice těžko implementuje a její výsledky jsou téměř neověřitelné. Proto se pro potřeby automatické predikce zaměříme na extrapolaci, neboli použití časových sérií. Za předpokladu, že se v datech v čase vyskytuje nějaká pravidelnost (a když ne dokonale pochopena, nebo podrobně změřena), naskytá se nám metoda, jež je obecně použitelná na situace, kdy jsou potřeba užitečná data k budoucímu vývoji. [4]

4.3.3 Trend

Trend odráží dlouhodobé změny v průměrném chování časové řady resp. obecnou tendenci vývoje zkoumaného jevu za dlouhé období. Je výsledkem faktorů, které dlouhodobě působí ve stejném směru, jako jsou například demografické podmínky či podmínky trhu v dané oblasti. Trend může mít různý charakter, může být rostoucí, klesající, strmý, mírný a v průběhu času se měnit, takže jej lze pokládat spíše za cyklus. Může být hladší než je vlastní časová řada, nebo také variabilnější. [1]

Jednou z možností, jak lze trend kvantifikovat, je model

$$\begin{aligned} X_t &= \alpha + \beta t + u_t \\ t &= 1, 2, \dots, T \end{aligned} \quad (1)$$

Který se označuje jako model lineárního deterministického trendu. Parametr β charakterizuje přírůstek řady X_t při změně času t o jednotku.

Dynamiku časové řady v jednotlivých obdobích lze také kvantifikovat pomocí tzv. měr dynamiky. Absolutní přírůstek (první difference) je definován jako

$$\begin{aligned} \Delta X_t &= X_t - X_{t-1} \\ t &= 1, 2, \dots, T \end{aligned} \quad (2)$$

a udává, jak se změnila hodnota v časové řadě v čase t ve srovnání s hodnotou v čase $t - 1$. Odečte-li se model lineárního trendu v čase $t - 1$, tj.

$$X_{t-1} = \alpha + \beta(t - 1) + u_{t-1} \quad (3)$$

od první rovnice pak získáme model ve tvaru

$$\begin{aligned} X_t - X_{t-1} &= \beta + e_t, \\ t &= 2, 3, \dots, T, \text{ kde } e_t = (u_t - u_{t-1}) \end{aligned} \quad (4)$$

Odhad parametru β se získá metodou nejmenších čtverců a má formu aritmetického průměru prvních diferencí. Je interpretován také jako průměrná diference (nebo průměrný absolutní přírůstek). [1]

Za předpokladu, že model trendu je exponenciální ve tvaru

$$\begin{aligned} X_t &= \gamma \delta^t \varepsilon_t \\ t &= 1, 2, \dots, T \end{aligned} \quad (5)$$

potom by bylo možné k jeho linearizaci použít logaritmickou transformaci, tj.

$$\ln X_t = \ln \gamma + \ln \delta t + \ln \varepsilon_t \quad (6)$$

Parametr $\ln \delta$ charakterizuje přírůstek řady $\ln X_t$ při změně času t o jednotku.

Čtvrtou možností, jak charakterizovat dynamiku řady, jsou koeficienty růstu a relativní přírůstky. Koeficient růstu je definován jako

$$\begin{aligned} k_t &= \frac{X_t}{X_{t-1}} \\ t &= 2, 3, \dots, T \end{aligned} \quad (7)$$

a po vynásobení stem říká na kolik procent hodnoty v čase $t - 1$ se změnila hodnota v čase t . V případě exponenciálního trendu časové řady je možné model v čase $t-1$ vyjádřit ve tvaru:

$$\begin{aligned} \frac{X_t}{X_{t-1}} &= \delta \frac{\varepsilon_t}{\varepsilon_{t-1}} \\ t &= 2, 3, \dots, T \end{aligned} \quad (8)$$

Po linearizaci logaritmickou transformací má tento model formu:

$$\ln X_t - \ln X_{t-1} = \ln \delta + \varepsilon_t$$
$$t = 2, 3, \dots, T, \text{ kde } \varepsilon_t = (\ln \varepsilon_t - \ln \varepsilon_{t-1}) \quad (9)$$

Odhad parametru $\ln \delta$ se získá metodou nejmenších čtverců a má formu aritmetického průměru prvních diferencí logaritmů časové řady. Po odlogaritmování se tento odhad interpretuje jako průměrný koeficient růstu a jedná se o geometrický průměr koeficientů růstu. [1]

4.3.4 Sezónnost

Sezónní složka zobrazuje pravidelně se opakující změny v časové řadě, které probíhají v jednom kalendářním roce a každý rok se opakují. Příčinami sezónních změn jsou nejčastěji vlivy změny ročních období nebo také různé společenské zvyklosti (výplata mezd v určité době, svátky, dovolené, apod.). [5]

Tato periodická oscilace, jak již víme, může mít dvě podoby odlišující se délkou vlny kolísání. V první řadě je to sezónní složka časových řad, kterou rozumíme periodicky se opakující obousměrné odchylky údajů řady od trendu, přičemž předpokládáme, že jde o odchylky opakující se s periodicitou jeden rok či kratší. V ekonomických časových řadách půjde nejčastěji o periodicitu měsíční nebo čtvrtletní. Oscilace vznikají ponejvíce v důsledku přímých či nepřímých příčin, které se rok co rok pravidelně opakují jako projev existence pravidelného střídání ročních období. [5]

Vedle toho vnímáme také oscilace s délkou vlny větší než jeden rok. Zde hovoříme o cyklické složce. Problematikou ekonomického cyklu se nebudeme zabývat, protože pro krátké období nemá na osobní finance zásadní vliv.

Předpokládáme-li reálnou existenci sezónní složky v časové řadě, přichází ke slovu nejprve kvantifikace sezónních výkyvů. Druhým důležitým úkolem pak bývá provést tzv. Sezónní očišťování. Jeho cílem je vyloučit sezónní složku z analyzované řady, protože sezónní kolísání do značné míry zakrývá základní dynamiku ekonomických jevů znemožňuje tím provádět kvalifikované srovnání vývoje v jednotlivých obdobích uvnitř roku. [1]

Modely sezónnosti mohou mít velice rozmanitou podobu. Záleží především na tom, jaké budeme formulovat předpoklady jednak o charakteru trendu analyzované časové řady, jednak

o vzájemném vztahu trendové a sezónní složky. Nejčastější cestou je buď technika konstrukce sezónních rozdílů a indexů, nebo klasický regresní přístup k sezónní složce. [5]

4.3.5 Korelace Dat

Korelace je statistická technika pro zjištění zda a případně kolik se dvě sady dat ovlivňují. Pro měření míry tohoto ovlivnění existuje mnoho metod, různých se přístupem i vstupy, které zvládají zpracovat.

Sledujeme-li pouze dvě proměnné, získáme první představu o jejich závislosti už pouhým uspořádáním zjištěných údajů do tzv. dvourozměrné tabulky. V hlavičce této tabulky se uvádějí hodnoty jedné proměnné, v legendě hodnoty proměnné druhé a v jednotlivých políčkách tabulky četnosti kombinací hodnot obou proměnných. Tyto četnosti nazveme sdružené četnosti a budeme je značit n_{ij} . V posledním řádku tabulky jsou sloupcové součty. Jsou tzv. okrajové (marginální) četnosti. Řádkové součty označíme $n_{i.}$ a sloupcové součty označíme $n_{.j}$. Symbolem n budeme značit rozsah souboru, symboly r a s počty řádků a sloupců tabulky (bez součtového řádku a součtového sloupce). Sledované proměnné, ať slovní nebo číselné, budeme značit x, y . Schématem dvourozměrné tabulky je pak tabulka č. 1.

Tab. 1: Schéma četnostní tabulky.

Hodnoty proměnné x	Hodnoty proměnné y				Součty čet- ností $n_{i.}$
	y_1	y_2	...	y_s	
$n_{i.}$	n_{11}	n_{12}	...	n_{1s}	$n_{1.}$
$n_{i.}$	n_{21}	n_{22}	...	n_{2s}	$n_{2.}$
...
$n_{i.}$	n_{r1}	n_{r2}	...	n_{rs}	$n_{r.}$
Součty čet- ností $n_{.j}$	$n_{.1}$	$n_{.2}$...	$n_{.s}$	n

Dvourozměrná tabulka se slovními proměnnými se nazývá kontingenční tabulka, ale s tou se v této práci nebudeme zabývat. Pro účely této práce je důležitá především tabulka numerických proměnných, která se nazývá korelační tabulka. V ní jsou v hlavičce uvedeny

bud' hodnoty proměnných (jde-li o nespojité proměnné nabývající jen malého počtu hodnot), nebo intervaly hodnot proměnných (jde-li o proměnné nabývající velkého počtu hodnot). Pokud se údaje korelační tabulky dále zpracovávají, jsou jednotlivé intervaly reprezentovány jejich středy. [5]

4.3.6 Bodový diagram

Bodový diagram se používá ke znázornění závislosti dvou číselných proměnných. Sledujeme-li u n určitých jednotek dvě číselné proměnné x a y , dostaneme celkem n dvojic hodnot (x_i, y_i) . Každá z těchto dvojic vyjádřena jako bod v pravoúhlé souřadnicové soustavě. Na horizontální osu se umísťuje stupnice hodnot vysvětlujících proměnné x a na svislou osu stupnice hodnot vysvětlované proměnné y . Roj vynesenech bodů pak informuje o charakteristických rysech závislosti obou proměnných. Lze z něj poznat především to, zda při růstu hodnot jedné proměnné mají hodnoty druhé proměnné tendenci růst či klesat, zda se tento růst či pokles zrychluje či zpomaluje apod. Lze z něj tedy poznat, jaký průběh má závislost obou číselných proměnných.

Z bodového diagramu lze získat i představu o těsnosti této závislosti. Proloží-li se (třeba od oka) rojem bodů čára vhodně popisující průběh závislosti, mohou jednotlivé body kolem této čáry kolísat někdy jen zcela nepatrně, někdy naopak značně. V prvním případě je možno považovat závislost za těsnou (silnou), ve druhém za slabou. Průběh a těsnost patří k nejčastěji sledovaným rysům závislosti dvou číselných proměnných.[5]

Přesto, že se jedná častou používanou metodu pro měření korelace dat, pro její grafickou povahu ji nelze příliš efektivně využít u automatizovaného systému, a proto ji v této práci budeme využívat z čistě ilustrativních důvodů.

4.3.7 Nelinearita

Problematika nelinearity je velmi široká a zdaleka ne prozkoumaná. Některé ekonomické časové řady jsou charakteristické strukturálními zlomy, změnami průběhu a variability. V této souvislosti se může v čase měnit i jejich autokorelační struktura. Tento způsob chování ekonomických časových řad nemůže být korektně zachycen lineárními modely. [1]

Nelinearita se u ekonomických časových řad může projevit odlišnými průměrnými diferencemi nebo průměrnými koeficienty růstu v různých obdobích.

4.4 Cash flow

Pojem cash flow má svůj původ v USA, kde si jeho zjišťování a sledování vynutily potřeby kapitálových trhů. Zpravidla se nepřekládá a znamená tok (angl. flow) peněz (cash). Od počátku minulého století se stává předmětem zájmu finanční teorie a v 60. letech se objevuje jako doporučení pro praxi rozšířit finanční výkazy podniků o výkaz cash flow. Vykazování cash flow se od roku 1971 stalo v USA (Opinion No. 19) povinnou součástí účetní závěrky podniků. Pozadu nezůstala ani Velká Británie, kde bylo zavedeno od roku 1975 povinné vykazování cash flow pro všechny podniky s ročním obratem nad 25 tisíc liber, ani Evropská unie, která ve 4. direktivě z roku 1978 doporučuje členským zemím sestavování tohoto výkazu. Výbor pro mezinárodní účetní standardy vydává v roce 1977 pod vlivem Opinion No. 19 mezinárodní účetní standard IAS 7, který upravuje jednotné vykazování cash flow.[citace] Tento historický vývoj jasně ukazuje na důležitost a hodnotu cash flow. Ovšem zatímco pro podniky je výkaz peněžních toků běžnou praxí, pro osobní použití jen pomalu získává oblibu. S rozmachem informačních technologií a integrováním informačních systémů do mobilních zařízení se spravování osobního plánu cash flow stává praktické i pro fyzické osoby. [6]

Cash flow je definován jako skutečný pohyb (tok) peněžních prostředků za určité období. Je východiskem pro řízení likvidity, neboť existuje rozdíl mezi pohybem hmotných prostředků a jejich peněžním vyjádřením (např. nákup na úvěr), vzniká časový nesoulad mezi hospodářskými operacemi vyvolávajícími náklady a jejich finančním zachycením, proto také vzniká rozdíl mezi náklady a výdaji a mezi výnosy a příjmy. [6]

V tomto případě jde o tzv. retrospektivní cash flow, která odpovídá na otázku, odkud se peníze a jejich ekvivalenty v daném časovém intervalu vzaly a kam se poděly. Uživatelé takto sledují absolutní i relativní intenzitu jednotlivých peněžních toků, posuny ve struktuře toků a míru jejich stability. Posuzují formy vázanosti kapitálu v aktivech podniku, způsoby jeho uvolňování a formy financování podniku.

Ve statickém pohledu představuje cash flow volnou zásobu peněz, které má subjekt k dispozici. V podobě ukazatele umožňuje měřit finanční situaci subjektu. [6]

V dynamickém pojetí jde o budoucí peněžní toky, které jsou součástí finančních plánů (rozpočtů) z pohledu investora o očekávaných příjmech z investice. [6]

4.4.1 Použitá struktura cash flow

Pro zobrazení a práci s cash flow je nutné stanovit nejmenší časovou jednotku pro výpočet. Nejčastěji se používá kalendářní měsíc. Protože je to velice přehledná jednotka a nezkušený uživatel s ní může jednoduše pracovat, budeme dále dělit peněžní toky na měsíční období. Dále je třeba peněžní toky rozdělit do jednoduše zpracovatelných skupin, těmto skupinám nadále říkáme příjmové a výdajové třídy. Příkladem výdajové třídy může být například oblečení, potraviny, nájem apod. Zaměstnanecký plat, nebo příjmy z podnikání by pak byly příjmové třídy.

4.4.2 Teoretický základ vykazování cash flow

U peněžních toků zvýšení aktiv v sobě váže potenciální snížení peněžních prostředků a naopak jejich snížení disponibilní peněžní prostředky uvolňuje. Peněžní toky se zadržují v aktivech (pohledávkách) i kapitálu (závazcích). Stojí na počátku a na konci obchodních transakcí subjektu a jejich pohyb přímo ovlivňuje peněžní toky. Pohledávky tvoří záchytný bod přeměny aktiv subjektu v peníze a vážou peněžní prostředky. Závazky (dluhy) podniku na druhé straně odkládají reálný úbytek peněz. Relaci mezi pohledávkami nebo dluhy podniku a cash flow vyjadřuje tabulka číslo 2.

Tab. 2: Vztah pohledávek nebo závazků ke cash flow.[6]

Změna stavu	Vliv na cash flow
zvýšení pohledávek	úbytek
snížení pohledávek	přírůstek
zvýšení závazků	přírůstek
snížení závazků	úbytek

4.4.3 Metody vykazování cash flow

V účetní teorii se obecně vyskytují dva odlišné přístupy ke zjišťování peněžních toků [6]:

- přímá metoda,
- nepřímá metoda.

Např. podniky využívají při sestavování výkazu cash flow oba přístupy současně, přímý i nepřímý. Pro zachycení pohybu peněz v osobních financích je vhodný ovšem pouze přímý způsob výpočtu cash flow. Oba přístupy si blíže rozebereme v následujících kapitolách.

4.4.4 Čistá přímá metoda

Je založena na sledování skutečných příjmů a výdajů a jejich agregaci do předem vymezených položek. Vychází přímo ze změn stavů peněžních prostředků nebo peněžních ekvivalentů, které jsou vyvolány finančně účinnými hospodářskými transakcemi. V tříbilančním systému jde o transakce ovlivňující současně výsledek hospodaření nebo o transakce působící změny rozvahy. Ačkoliv se jeví teoreticky tato metoda sestavení výkazu cash flow jako bezproblémová, bankovní operace, zachycené na příslušných účtech, neumožňují zjistit účel, za kterým byly provedeny. Sledování skutečných toků peněz a jejich ekvivalentů podle jednotlivých druhů by vyžadovalo zavést nové syntetické účty příjmů a výdajů, které by se uzavíraly prostřednictvím bilance cash flow a jejich saldo by se přeneslo na účty peněžních prostředků a peněžních ekvivalentů. [6]

4.4.5 Nepřímá metoda

Nazývá se takto proto, že vychází nikoliv z výnosů a nákladů, ale z docíleného výsledku hospodaření, který dále upravujeme. Tato metoda vychází z výkazu zisku a ztrát podniku, přesněji řečeno z výsledného salda mezi výnosy a náklady, které transformujeme na cash flow. Lze ji použít pouze u peněžních toků z hlavní výdělečné činnosti subjektu, a to jen na tu část peněžních toků, která se nevykazuje jako hrubé peněžní tok. Spočívá v úpravě zisku či ztráty z hospodaření o nepeněžní položky a o změny položek rozvahy, vyjadřujících rozdíl mezi toky příjmů a výdajů a mezi toky výnosů a nákladů.

Nepřímá metoda je obecně považována za snažší a nenáročnou na vstupy. Všímá si pouze těch peněžních toků z provozní činnosti, které jsou vykázány na netto bázi, a neobsahuje tedy žádné platby. Pro subjekt je výhodná i proto, že externímu uživateli vyzradí o charakteru a struktuře peněžních toků méně, než metoda přímá. [6]

Z předchozích odstavců je zřejmé, že tato metoda je naprosto nevhodná pro použití v osobním plánování. Pro osobní finance nemá výkaz zisků a ztrát hodnotu a proto s ním pro osobní potřebu nikdo nedisponuje.

4.5 Forma predikce

Aby byly data lehce matematicky zpracovatelná, je nutné jasně stanovit jejich formu. Proto si definujeme následující zásady.

- Data budou rozdělena podle druhu. Tyto druhy mohou být stanoveny uživatelem a budou sdružovat položky, jejichž predikci si přeje.
- Výstupem predikce budou předpokládané měsíční obraty za daný druh.
- Základní časové období pro vyjadřování dat z minulosti bude kalendářní měsíc.

4.6 Měření přesnosti predikce

Pro měření přesnosti predikce použijeme v této práci vlastní ukazatel, jež nadále budeme nazývat míra přesnosti. Ta bude udávat, na kolik shoduje predikovaná hodnota s tou skutečnou. Ukazatel bude nabývat hodnoty od nuly do jedné respektive od nuly do sta procent. Pokud se predikovaná hodnota liší o dvojnásobek reálné hodnoty, je přesnost podle ukazatele nulová. Pro vyjádření souhrnné přesnosti za delší období, bude použit aritmetický průměr za jednotlivé měsíce. Geometrický průměr u tohoto ukazatele bohužel nelze použít kvůli měsícům, kdy je míra přesnosti nulová.

4.7 Výdajové a příjmové typy

Pro další potřeby práce je třeba rozdělit příjmy a výdaje na jednotlivé typy podle způsobu jejich vlivu na osobní bilanci.

- **Pravidelné položky neměnné.**
 - Pravidelné příjmy neměnné: nejtypičtějším zástupcem příjmů tohoto typu je zaměstnanecký plat bez pohyblivé složky. Tento typ příjmu se mění skokově při zvýšení či snížení platu. Jinak zůstává dlouho stabilní. Dále se může jednat o výplatu starobního či invalidního důchodu, sociálních dávek, apod.
 - Pravidelné výdaje neměnné: nejčastěji se jedná o nájem, zálohy, splátky dluhů či pojistek a jiné cyklické platby, které se mění jen velmi zřídka.

○ **Pravidelné položky pohyblivé.**

- Pravidelné příjmy pohyblivé: jedná se například o plat s pohyblivou složkou, či cyklické příjmy z podnikání. Jeho výše se neustále mění a může vykazovat výrazné sezónní výkyvy.
- Pravidelné výdaje pohyblivé: tento typ výdajů je běžnou položkou v osobním rozpočtu, ale pouze málokdy dosahuje stejných hodnot. Typickým příkladem jsou například potraviny, útraty v restauracích, pohonné hmoty nebo telefonní účty.

○ **Nepravidelné položky.**

- Nepravidelné příjmy: do tohoto typu můžou spadat nepravidelné příjmy z podnikání, výnosy z prodeje majetku, dědictví a všechny ostatní výnosy, jež se cyklicky neopakují, či nevykazují žádnou pravidelnost.
- Nepravidelné výdaje: jedná se o typ výdajů, který běžně neprovádíme. Často sem spadají nákupy drahých položek, ale také může jít o levnější platby, které se cyklicky neopakují. Může sem zařadit například opravy, nákupy nemovitostí, dopravních prostředků nebo vybavení domácnosti.

4.8 Metodika návrhu aplikace

Pro návrh aplikací se dnes využívají v podstatě dva způsoby vývoje. Klasický (tj. kaskádový nebo také vodopádový) nebo iterační. Klasický způsob je původní způsob vývoje softwaru a probíhá tak, že se nejprve provede analýza a návrh, poté se aplikace naprogramuje a otestuje. Nakonec je systém nasazen a udržován. Vše probíhá s důrazem na to, že jednotlivé části se neprolínají a jedna navazuje na druhou. Tento model tedy vyžaduje, aby se k následující fázi přikročilo pouze tehdy, pokud je ta předcházející kompletní a perfektně připravená. Nejprve se například připraví specifikace požadavků, které jsou pevně dané. Teprve když jsou požadavky úplně kompletní, přejde se k návrhu. Požadovaný software je navržen a návrh předán implementátorům – návrh by měl být jakýmsi plánem implementace daných požadavků. Když je tedy návrh hotový, programátoři se ujmou jeho implementace. V pozdějších etapách této implementační fáze jsou vytvořené softwarové komponenty kombinovány, zavádí se nová funkcionality a odstraňují chyby.[13]

Dnešní moderní programování ovšem vychází z principu, že při psaní mohou vzniknout chyby (a často taky vznikají), a dále z toho, že některé myšlenky je vhodné si ověřit v praxi, než se zahrnou do projektu. Proto se využívá postup, při kterém se kroky analýza/návrh, programování a testování mohou provádět současně, nebo se v některých fázích vracet k předchozím. S důrazem na to, že tyto tři části mohou probíhat současně nebo se v některé fázi vracíme k předchozí. Takovýto způsobu říkáme iterační a v ten také bude použit v této práci. Například programátor si chce ověřit funkčnost návrhu, a proto vytvoří tzv. „minimální implementaci“ dané části, to znamená, že programování předbíhá návrhu. Nebo se při testování zjistí, že datová prostupnost aplikace je nižší než požadovaná, a proto se daná část pře-programuje. V ideálním případě se testování účastní i zákazník, který zhotovení programu zadal.

5 Automatizace predikce cash flow

5.1 Rozložení cash flow na jednotlivé kategorie

Jak již bylo uvedeno výše, bude se pro predikci provádět extrapolace jednotlivých výdajových a příjmových kategorií a z těch poté bude sestaven celkový pohled na tok peněžních prostředků za určité období. Tyto kategorie budou sdružovat logicky související peněžní toky a budou volitelné uživatelem aplikace. Predikce takto bude mít pro uživatele daleko větší hodnotu, protože bude mít jasně vyjádřeno, kolik prostředků získává, nebo vydává za jednotlivé oblasti svého života. Po pochopení struktury svých financí může poté např. efektivněji zvolit, na čem bude šetřit. Po provedení extrapolace těchto kategorií bude sestaven i celkový peněžní tok, z kterého poté uživatel vyčte, zda vyjde se svým rozpočtem.

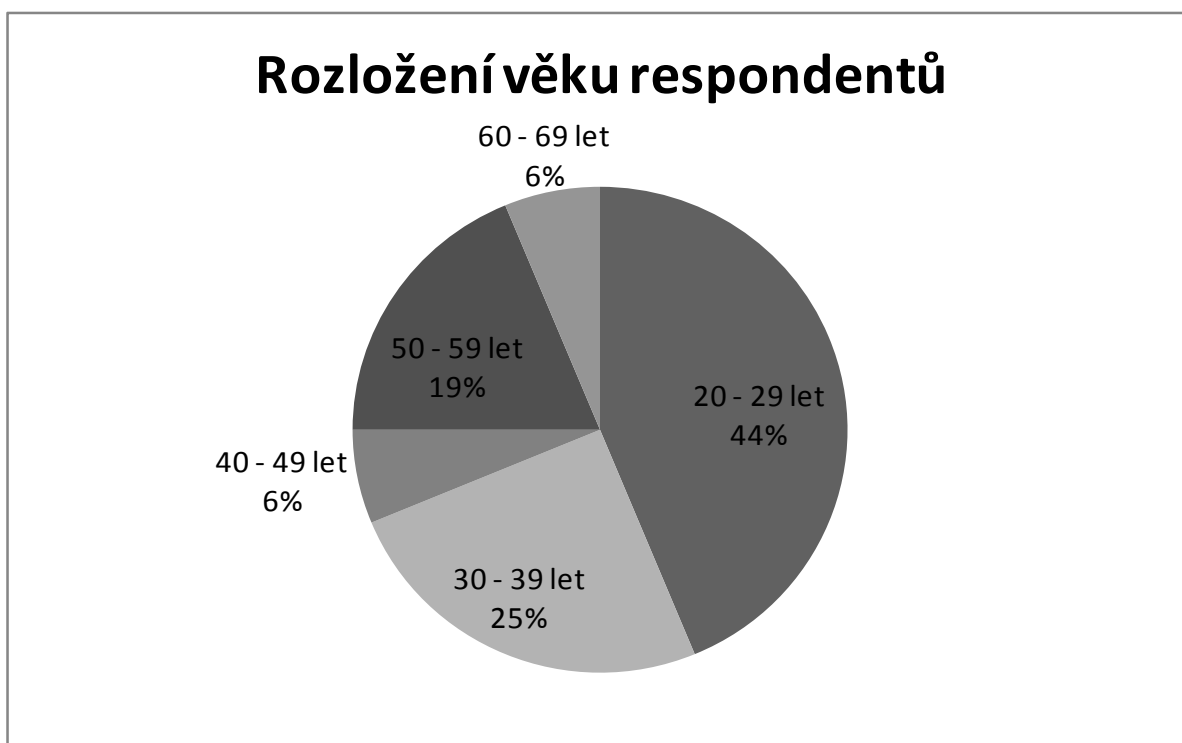
5.2 Data pro testování algoritmu a složení vzorku

Pro testování výsledků předpovědi bylo potřeba získat dostatek časových řad, které vyjadřují měsíční pohyby respondentů. Za tímto účelem byla provedena anketa, jež měla za úkol tyto řady získat. Takovýto výzkum je velice složitý, protože většina osob nechce zveřejnit přesnou strukturu svých příjmů a výdajů, protože je považuje za velice soukromá. Respondenti proto byli ujištěni, že jsou data naprosto důvěrná a budou použity jen pro tento výzkum a že nebudou evidovány jakékoliv informace, jež by mohly prozradit jejich identitu. Dalším problémem bylo, že respondenti musejí již mít agendu svých příjmů, aby řady měly dostatečnou délku, na níž lze provádět analýzu. Ve výsledku byly do ankety přijaty data od 16-ti respondentů. V průměru měl každý respondent rozděleny svoje pohyby do 9,1 kategorií. Pro testování bylo proto k dispozici 146 časových řad různých délek. Respondenti vyplnili také krátký dotazník, ve kterém uvedli svůj věk, jakým způsobem vedou svůj peněžní deník a poté měli za úkol uvést na příštích 6 měsíců odhad svých příjmů a výdajů pro každou kategorii. Tento prvotní průzkum byl proveden v dubnu 2011. Všichni respondenti byli poté požádáni, aby příštího půl roku pečlivě vedli svůj finanční deník svou obvyklou metodou. V listopadu 2011 byly poté sesbírány od těchto respondentů peněžní deníky za uběhlé období. V tomto druhém kole průzkumu se podařilo získat data od 14-ti z původních 16-ti respondentů. Respondenti projevovali o výzkum zájem a většina z nich potvrdila, že se svou stávající metodou evidence nejsou spokojeni a že vidí přínos jakékoliv aplikace, která pomůže plánování jejich rozpočtů.

Forma, jakou respondenti předali data o svých pohybech, nebyla přesně definována. Čtyři respondenti proto předali data na papíře a zbytek je odevzdal ve formátu xls aplikace Microsoft Office Excel.

5.2.1 Respondenti

Respondenti jsou ve věku od 21 do 63 let. Z grafu č. 2, kde vidíme věkové rozložení respondentů je zřejmé, že anketa získala data především od mladých lidí mezi 20-ti a 30-ti lety. Objem studentů ovšem tvořil méně než 19% dotázaných a výsledky jsou proto především od lidí s pravidelným příjmem a tím pádem data dobře reprezentují chování osob v produktivním věku (pro které je aplikace určena především).



Obr 2: Rozložení věku respondentů ankety.

Drtivá většina respondentů žije v Brně a jeho okolí. Vzhledem k tomu, že se ale nakládání s osobními financemi výrazně neliší podle místa bydliště, neměl by tento fakt ovlivňovat výsledky výzkumu.

5.2.2 Časové řady

Respondenti měli za úkol odevzdat údaje o měsíčních výdajích a příjmech rozdělených podle kategorií, tak jak jim přijde pohodlné. Všechny pohyby byly vyjádřeny v českých korunách.

Zároveň obdrželi instrukce jak řady očistit o zvláštní pohyby, které se periodicky neopakují. V tomto případě se jednalo hlavně o nákupy vozidel, nemovitostí a zařízení domácnosti. Zároveň obdrželi vysvětlení k tomu, co bude aplikace provádět, a instrukci označit kategorie tak, aby bylo zřejmé, zda by v aplikaci u dané kategorii aktivovali automatickou prognózu, či nikoliv.

Po ukončení sběru ankety, bylo k dispozici 146 řad různých délek. Z toho 28 řad bylo příjmových a 118 výdajových. Nejdelší z nich dosahovaly délky 74 měsíců. Respondenti zaznamenali v anketě, že by aktivovali automatické prognózování u 81 % řad.

5.3 Datové případy, jejich definice a extrapolace

V této sekci si definujeme datové případy, které bude aplikace sledovat a v závislosti na jejich typu poté provádět danou extrapolaci jednotlivých kategorií. Pro každý datový případ definujeme vstupní požadavky, navrhujeme postup detekce a poté postup extrapolace na příštích 12 období. Pro potřeby modelování není důležité, zda se bude jednat o příjmy nebo výdaje, a proto budeme uvažovat pouze jeden typ pohybů. Zkoumáme objem peněz který je každý měsíc zaregistrován. Pro názornost si zobrazíme, jak takováto řada může vypadat a jak by se měl algoritmus zachovat v případě její predikce.

5.3.1 Datový případ – silná sezónnost

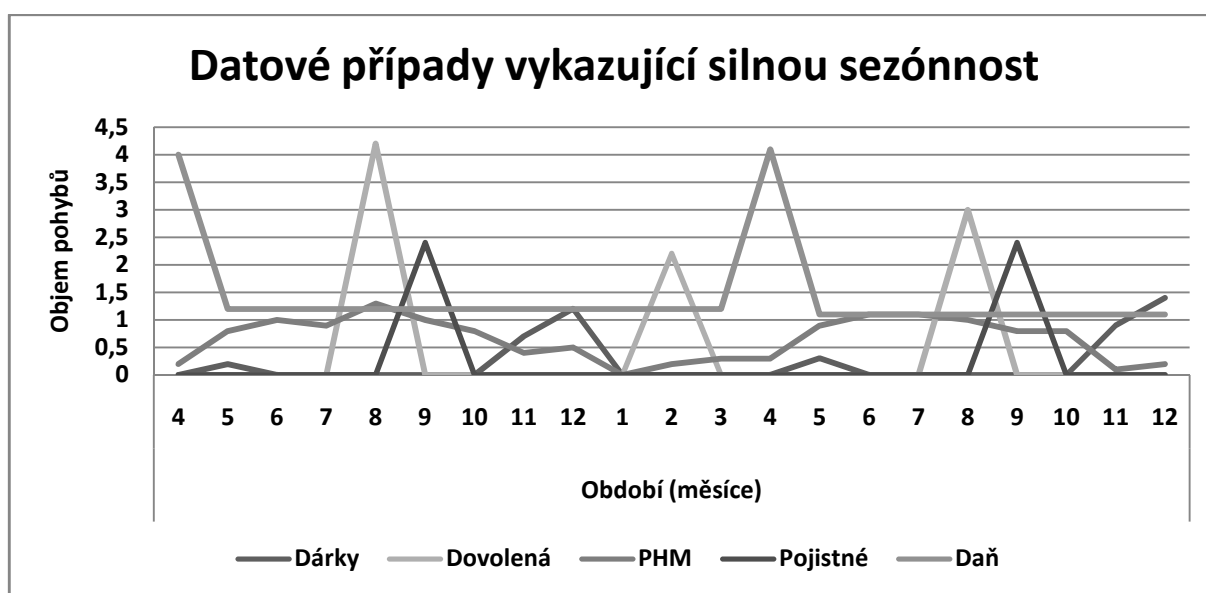
Chování časové řady je silně ovlivňováno sezónními vlivy působící na daný peněžní pohyb. Tento případ je jedním z neklíčovějších při projekci dat. Tento případ zachycuje celé spektrum příjmů i výdajů. Příčinami těchto sezónních změn jsou nejčastěji vlivy změny ročních období nebo také různé společenské zvyklosti a právě díky těmto vlivům celá řada kategorií vykazuje velkou míru sezónnosti. Perioda sezónnosti se může pro každou položku lišit, ale nejčastěji se setkáme s periodou roční. Jsou ovšem i specifické pohyby v osobních financích vykazující jiné periody, nečastěji se jedná ještě o periodu tříměsíční a šestiměsíční.

Tab. 3: Příklad silné sezónnosti, údaje jsou v tis. Kč.

Kategorie	Období (měsíce)																		
	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Dárky	0	0	0	0	0	0,7	1,2	0	0	0	0	0,3	0	0	0	0	0	0,9	1,4
Dovolená	0	0	4,2	0	0	0	0	0	2,2	0	0	0	0	0	3	0	0	0	0
PHM	1	0,9	1,3	1	0,8	0,4	0,5	0	0,2	0,3	0,3	0,9	1,1	1,1	1	0,8	0,8	0,1	0,2
Pojistné	0	0	0	2,4	0	0	0	0	0	0	0	0	0	0	0	2,4	0	0	0
Daň	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	4,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1

V tabulce č. 3 vidíme několik případů z finančních deníků, které vykazují velice silnou sezónní závislost. Jsou to například kategorie, jež jsou nějakým způsobem vázány na svátky, či slavnosti. Například nákupy dárků, budou vždy četnější v období vánoc a v době narozenin blízkých. Na dovolenou lidé také většinou vyráží každý rok v podobnou dobu. Některé položky v osobních financích mohou být vázány také na zvláštní osobní zvyklosti. V našem příkladu je například osoba, která používá automobil převážně k dojíždění na svou rekreační chatu, kterou ale navštěvuje převážně v létě a v zimě automobil využívá jen zřídka. U takové osoby budou i výdaje za pohonné hmoty vykazovat silnou sezónnost.

Zvláštním případem sezónnosti jsou platby probíhající jednou ročně. Zde to mohou být platby nejruznějších pojištění, daní, nejruznějších služeb, vklady do spořicích nástrojů apod. Tento typ výdajů bývá vázán na konkrétní měsíc a proto je jeho identifikace relativně snadná a predikce může být velice přesná.



Obr 3: Datové případy vykazující silnou sezónnost.

Pro extrapolaci dat se sezónní složkou je ovšem nutné mít dostatek dat za mnoho období do minulosti. Toto v případě internetové aplikace, která sbírá data uživatelů postupně, může být velkým problémem. Pokud uživatel začne aplikaci používat, a nezadá data z minulých let, bude aplikaci trvat více než dva roky, než uživateli dokáže systém nabídnout tento typ predikce. I v tomto případě půjde ovšem o velice hrubý odhad, a přesto že u některých typů pohybů bude dosaženo dobrých výsledků, v jiných kategoriích může být extrapolace na základě takto krátké datové sady dosti zavádějící. Navíc je nutné podchytit trend, měnící se v průběhu života osoby. Ta s měnícími se životními fázemi může měnit výše vydávané za jednotlivé kategorie, přestože míra sezónnosti bude zůstat stejná. Tento trend by měla aplikace zjistit a promítnout do prognózy.

Aplikace by měla dopady nevýhod tohoto datového případu minimalizovat. Zprv by měla zkrátit dobu, za jakou dokáže odhalit sezónnost v datech a zadruhé by měla upozorňovat uživatele při předpovědi s malým množstvím historických dat, aby si byl vědom toho, že predikce nemusí být příliš přesná.

- V době, kdy ještě není možné provádět zkoumání sezónnosti (s historickými daty do dvou let), by aplikace měla prozkoumat možný výskyt sezónnosti s periodou kratší než jeden rok a použít podle toho odpovídající extrapolaci. Poté co data dosáhnou dostatečné délky, může aplikace přejít k predikci sezónnosti s periodou jednoho roku.
- Poté kdy data v aplikaci dosáhnou dostatečné délky, aby šla provést predikce s aplikací sezónních vlivů, může aplikace přejít k extrapolaci s pomocí této metody. Pokud je ovšem délka dat na spodní hranici, odkdy lze tuto metodu aplikovat, měl by být uživatel odpovídajícím způsobem upozorněn. Aplikace by měla uživateli nějakým způsobem sdělit, že data zatím nedosahují příliš velké délky a predikce tak může podávat zkreslená data. Ten se poté sám na základě predikce může rozhodnout, zda extrapolace dodává užitečná data a v opačném případě může predikci pro danou kategorii vypnout.

Pokud hovoříme o sezónnosti, rozumíme tím především sezónnost s roční periodou. V případech, kdy ovšem nemáme dostatečně dlouhá data pro spolehlivé testování roční sezónnosti, můžeme pak testovat data na sezónnost s kratší periodou. U osobních financí se můžeme setkat ještě s 6-ti měsíční a 3 měsíční periodou pohybů. Takovéto případy můžou nastát

vat například při platbách za pojištění apod. Sezónnost tedy rozdělíme na roční, půlroční a čtvrtletní, z nichž poslední dvě se budou používat pouze při nedostatku dat.

Pro čtvrtletní sezónnost musejí data dosahovat minimální délka dat 6 měsíců, pro pololetní poté 12 měsíců a pro roční sezónnost 24 měsíců. Zároveň je třeba časové řady očistit o nepravidelné výdaje, které výrazně vybočují z běžného chování spotřebitele, jež mohou úspěšnou detekci a extrapolaci takovýchto časových řad znemožnit.

Charakter projekce sezónnosti je takový, že odhalí i sezónnost s kratší periodou. Proto není třeba provádět např. čtvrtletní predikci, pokud již máme dostatečná data pro roční sezónnost.

Odhalit, zda časová řada vykazuje sezónní chování, není jednoduché ani pro zkušeného analytika natož pro automatizovaný systém. Proto byl vyvinut vlastní algoritmus, který je vhodný pro automatizované stanovení míry sezónnosti. Postup výpočtu této míry si zde nyní rozebereme. Postup bude popsán pro délku periody 12 období, ale tento postup lze aplikovat i na čtvrtletní a pololetní sezónnost.

Pro výpočet míry sezónnosti je nejprve nutné určit, jak se v každém měsíci částka odlišuje od ročního průměru. Tento roční průměr si označíme a jednotlivé měsíční částky pak n , dolními indexy r a m budeme značit rok respektive měsíc, ke kterému údaj přísluší.

$$a_r = \frac{\sum_{m=1}^{12} n_{rm}}{12} \quad (10)$$

Nejprve tedy spočítáme průměrnou měsíční částku pohybů v daném roce. Tento postup nám zajistí měření odchylky nezávisle na trendu, jež může nastávat v životě jedince, jak bylo popsáno výše. Pokud si koeficient měsíční odchylky označíme k , pak:

$$k_{rm} = \frac{n_{rm}}{a_r} \quad (11)$$

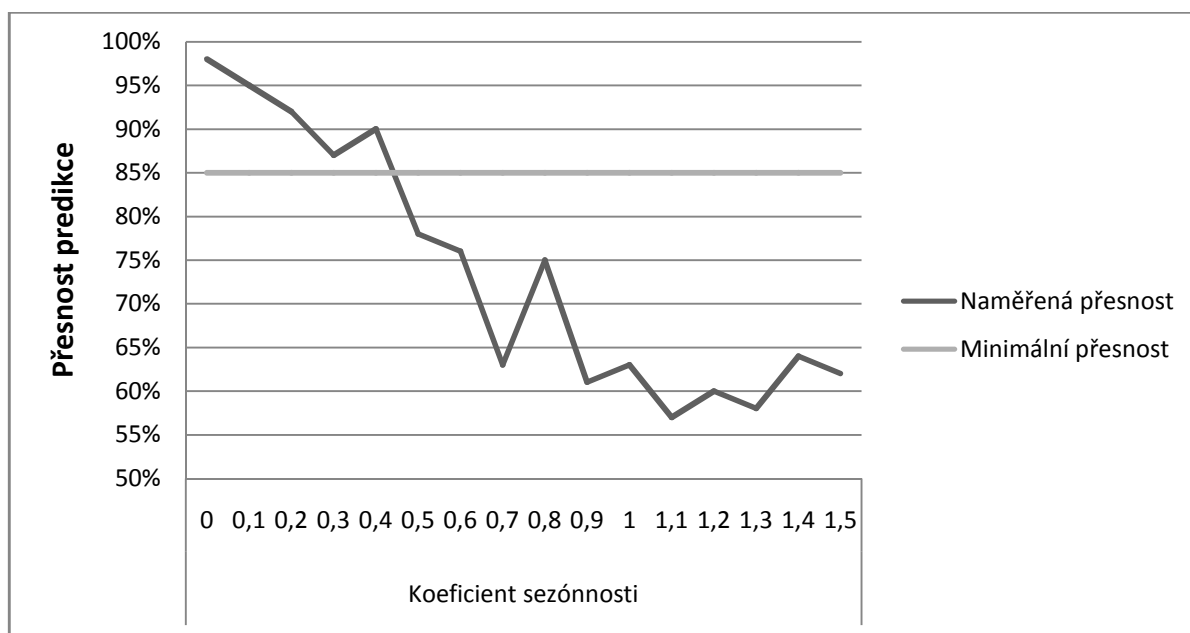
Díky tomu, že měříme odchylku od průměru pomocí poměru a ne rozdílem, bude koeficient správně reagovat na měsíce, v nichž není žádný pohyb. V takovém případě bude koeficient nulový a tím pádem se bude dobře porovnávat oproti ostatním obdobím. Poté co si vyjádříme koeficient měsíční odchylky, zjistíme průměrnou odchylku tohoto koeficientu každého měsíce. Tuto odchylku si označme o . Ve výpočtu použijeme také průměrnou hodnotu koeficientu měsíční odchylky, kterou označíme a_{krm} . Výpočet poté vypadá následovně.

$$O_m = \frac{\sum_{r=1}^{\text{počet let}} |k_{rm} - a_{krm}|}{\text{počet let}} \quad (12)$$

V této fázi výpočtu již znám míru sezónnosti, jež vykazují jednotlivé kalendářní měsíce. V této fázi zbývá určit z těchto čísel roční koeficient sezónnosti, označme si jej S . Pro jeho výpočet se logicky vybízí aritmetický průměr, jež se běžně používá při hodnocení koeficientů. Takovýto výpočet ovšem v našem případě není vhodný. Jednotlivé míry sezónnosti budou často nabývat nulové hodnoty, symbolizující nejvyšší míru sezónnosti. Tyto hodnoty jsou proto velice důležité. Pro výpočet musíme proto využít aritmetický průměr, který lépe odráží míru sezónnosti v našem případě. Výpočet tohoto koeficientu poté vypadá následovně:

$$S = \frac{\sum_{m=1}^{12} O_m}{12} \quad (13)$$

Poté co zjistíme míru koeficientu sezónnosti, musíme si určit kdy prohlásit, že míra sezónnosti je dostatečně vysoká pro predikci. Měřením na testovacích datech byly zjištěny následující míry přesnosti pro koeficient sezónnosti, při zapojení predikce pomocí sezónní extrapolace.



Obr. 4: Přesnost predikce v závislosti na naměřené míře sezónnosti historických dat

Z grafu na obrázku č. 4 vidíme klesající přesnost sezónní předpovědi v závislosti na naměřené míře sezónnosti dat. Koeficient sezónnosti vykazuje absolutní sezónnost v bodě 0. To znamená, že pro dosažení nulové hodnoty musejí být prostředky rozloženy každý rok úplně stejně. Zvyšující se koeficient sezónnosti signalizuje odchylku od dokonale sezónního chová-

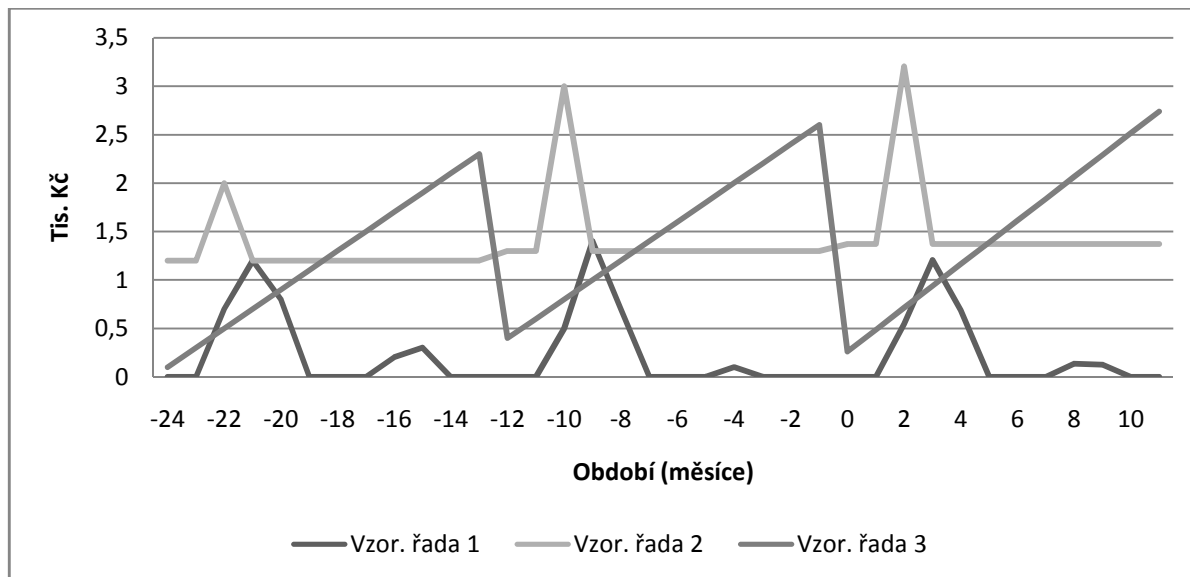
ní. Pro stanovení míry sezónnosti bylo provedeno měření nad testovacími daty při různých měřích koeficientu sezónnosti a to jedné desetině od 0 do 1,5. Pokud si stanovíme, že požadujeme alespoň osmdesátipětiprocentní přesnost předpovědi, bude jí dosahováno při hodnotě koeficientu nižší než 0,4.

Z tohoto měření byla proto stanovena podmínka, která říká, že při hodnotě koeficientu sezónnosti pod hodnotou 0,4, můžeme prohlásit o časové řadě, že vykazuje sezónní chování a pro její extrapolaci můžeme použít metodu sezónní extrapolace.

Extrapolace dat, které vykazují silnou sezónnost, bude probíhat v následujících krocích. Pro potřeby výpočtu si nejprve stanovíme koeficienty odchylky pro každý měsíc a rok. Tento koeficient si budeme nadále značit k_{rm} a jeho výpočet bude vypadat následovně:

$$k_{rm} = \frac{n_{rm}}{\left(\frac{\sum_{m=1}^{12} n_{rm}}{12} \right)} \quad (14)$$

Z těchto koeficientů bude poté stanoven průměr za všechny roky a bude jím vynásobena průměrná částka, jež minulý rok tato řada vykazovala. Tímto způsobem dostaneme prognózu dat s ohledem na sezónnost časové řady.



Obr. 5: Ukázka prognózy na vzorových datech vykazujících silnou sezónnost.

Graf na obrázku č. 5 ilustruje, jak bude vypadat prognóza pro vzorová data. Období 0 představuje moment provedení prognózy. Kladné období poté představují prognózu.

5.3.2 Datový případ – stejná data a schodky

Dalším častým případem jsou příjmy a výdaje, které se meziměsíčně nijak nemění a dosahují neustále stejných hodnot. Takováto charakteristika časové řady může být vyvolána celou řadou běžných peněžních pohybů jedince. Jedná se všechny stálé výdaje a příjmy s měsíční periodou. Do tohoto datového případu spadá například nájem, splátky dluhu, paušální platby za služby apod. Nejvýznamnější časová řada, která často vykazuje takovouto charakteristiku, je ovšem příjem ze závislé činnosti.

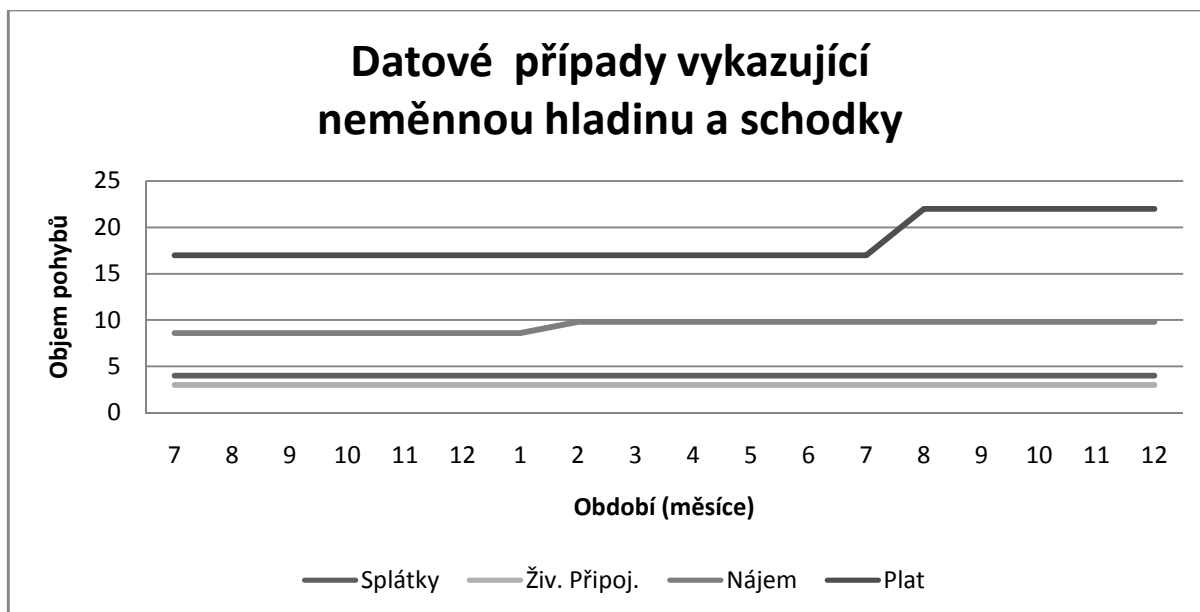
Tab. 4: Příklady neměnných dat a schodků řad

Kategorie	Období (měsíce)															
	$t-16$	$t-15$	$t-14$	$t-13$	$t-12$	$t-11$	$t-10$	$t-9$	$t-8$	$t-7$	$t-6$	$t-5$	$t-4$	$t-3$	$t-2$	t
	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Splátky	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Živ. Připoj.	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Nájem	8,6	8,6	8,6	8,6	8,6	9,8	9,8	9,8	9,8	9,8	9,8	9,8	9,8	9,8	9,8	9,8
Plat	17	17	17	17	17	17	17	17	17	17	17	22	22	22	22	22

Tyto řady se vyznačují dlouhými obdobími, kdy se výše částky každý měsíc opakuje. V takovéto řadě ale poté může nastat změna, nazýváme ji nadále schodek. Po této změně se bude dále opakovat nová částka po změně. Takovéto schodkovité chování je vyvoláno změnou výše pohybu a nejčastěji se jedná o změnu ceny placené služby. V případě příjmu jde pak o zvýšení platu, či změnu zaměstnání.

Zvláštním případem je schodek do nuly, svědčící o tom, že uživatel již nadále službu nepoužívá, či nepobírá plat. Tomuto případu se budeme věnovat v oddílu speciálních datových případů.

Nepříjemnosti v těchto situacích vyvolávají pohyby, jež jsou charakterizovány placením záloh a poté doplácením či vratkami z těchto záloh. Ty vnesou do neměnné řady výkyv, jež narušuje souvislost dat a jsou těžko obecně popsatelné. Tím pak mohou znemožnit správnou klasifikaci takovéto časové řady a její správnou extrapolaci. Uživatel by měl být proto aplikací poučen, aby doplňoval pohyby tohoto typu do zvláštní sekce. Například jednu kategorii pro zálohy na daň a jinou poté na doplatky či vratky. Tak lze zaručit jednodušší detekci a přesnější extrapolaci takovýchto dat.



Obr. 6: Příklady neměnných řad a schodků v nich.

Tyto řady mají tu výhodu, že jdou nasadit i u časových řad, jež mají i velmi malou délku. Pokud se částka za poslední dva měsíce zopakovala, je pravděpodobné, že i tento a další měsíce bude následovat ve stejné výši. Definujme si proto, že již tři po sobě následující období znamenají neměnnou řadu s tím, že nesmí více než jednou ročně projevit schodek této hladiny.

Abysme mohli o časové řadě prohlásit, že spadá do tohoto datového případu, musíme nejdříve disponovat alespoň daty za tři období, data navíc za rok musí nabývat nejvíce dvou různých hodnot. Zároveň musíme data očistit o případné výkyvy způsobené nepravidelnými výdaji, které výrazně vybočují z běžného chování spotřebitele.

Poté co aplikace na základě stanovených podmínek zjistí, že se jedná o tento datový případ, zkopíruje pouze objem pohybů z minulého měsíce. Tak nabídne uživateli ulehčení při práci s jeho finančním plánem.

5.3.3 Datový případ – nepoužívaná řada

Jedná se příjem či výdaj, jež v minulosti měl význam, nicméně pohyby tohoto typu se již nadále v peněžním deníku nevyskytují. Algoritmus musí detekovat tuto charakteristiku v datech a nadále prognózovat pro tuto časovou řadu nulové pohyby.

Vstupní požadavky pro tento datový případ nejsou žádné a detekce takovéto řady je velice jednoduchá. Pokud časová řada neobsahuje za posledních 12 měsíců žádná data, je na místě domnívat se, že se s danou kategorií nikdy nepracovalo a nebo již nepracuje. Prognóza pro

takovéto řady se poté vůbec neprovádí. Uživatel by si ovšem měl být vědom toho, že s danou kategorií nadále nechce pracovat a její predikci ihned vypnout. Jinak bude trvat 12 měsíců, než algoritmus sám tuto řadu z predikování vyřadí. Uživatel by měl být o tomto chování dobře informován.

Nevýhodou tohoto přístupu je, že nelze provádět predikci pro pohyby s periodou větší než jeden rok. V lidském životě jsou takovéto opakující se příjmy či výdaje velice vzácné, a pokud se vyskytují, jejich nepravidelnost většinou stejně neumožňuje spolehlivou predikci.

5.3.4 Datový případ – velice krátká řada

Velice specifickým případem jsou krátké řady. Pokud časová řada obsahuje pouze údaje za aktuální měsíc, či za velice krátké období, nelze zde uvažovat o žádné pokročilé metodě prediktivní analýzy. Nicméně aplikace může uživateli i přesto nabídnout užitečné informace. V některých případech, kdy je charakter řady silně lineární, může aplikace dokonce provést přesnou prognózu. Při takovýchto objemech dat je ovšem nutné brát predikci s velkou rezervou a uživatele výrazně upozornit, že se jedná pouze o velice nepřesný odhad a že se bude prognóza s přibývajícími daty zpřesňovat. Jak si níže vysvětlíme, tak u tohoto datového případu je také velice důležité poskytnout uživateli nějaký způsob zpětné vazby.

Tab. 5: Charakteristiky velice krátkých řad.

Charakteristika	Období (měsíce)				
	$t-4$	$t-3$	$t-2$	$t-1$	t
Pouze aktuální období	0	0	0	0	4
Pouze období $t-1$	0	0	0	2	0
Pouze jedno období	0	0	4	0	0
Pouze minulé s aktuálním	0	0	0	2	2,4

V tabulce 5 můžeme vidět všechny situace, které by měla aplikace rozlišit jako zvláštní případy.

- Jedná se o nově vytvořenou kategorii a byly vloženy data pouze za aktuální měsíc. Jak si níže vysvětlíme, je prognóza v této fázi spíše ukázkou funkce aplikace, než predikcí. V případě jediného datového bodu máme jen dvě reálné možnosti a to rozkopírovat tuto částku do budoucích měsíců, či predikci neprovádět. Z důvodů níže uvedených zvolíme možnost rozkopírování. Jedná se o jediný případ, který

využívá data z aktuálního měsíce, všechny ostatní metody tento měsíc ignorují, jako nekompletní.

- Další možností je, že je vyplněn pouze měsíc minulý, či minulý s aktuálním. V tomto případě předpokládáme, že aktuální měsíc je ještě nekompletní. Situace je obdobná jako v prvním případě a proto rozkopírujeme pouze měsíc minulý.
- Posledním případem je situace, kdy je vyplněno pouze jedno období a nejedná se o poslední období. Dá se předpokládat, že se jedná o ojedinělý pohyb a nelze proto říci, kdy se bude opakovat.

Jednou nebo více z těchto fází projde každá kategorie, jež nebude vyplněna za pomoci historických dat. Proto je zde třeba pozorně nastavit chování aplikace tak, aby byl uživatel motivován nadále používat systém. Uživatel, který prvně pracuje s aplikací a zadal prvních několik dat, musí dostat zpětnou vazbu na svou činnost a musí pochopit jak aplikaci používat a v čem mu může pomoci. Proto přestože v těchto fázích spíše o hádání, než prognózu, je důležité uživateli nabídnout data. V této fázi je ovšem naprosto nezbytné popsat uživateli charakter predikce a vysvětlit mu, že se prognóza bude s přibývajícími daty měnit.

5.3.5 Datový případ – řada bez historických dat

Jedná se o řadu, která neobsahuje žádná data z historie. Tímto stádiem musí projít každá kategorie nejméně jednou a to ihned po vytvoření, před přidáním jakýchkoliv dat. V této situaci je samozřejmě pouze jedna možnost a to prognózu vůbec neprovádět. Některé návrhy algoritmu by tuto situaci řešili automaticky, např. tím že by spočítali průměr z nulových dat, ovšem z důvodu šetření systémových prostředků je dobré tuto situaci podchytit a vyhnout se tak průchodu celým algoritmem.

Zajímavým případem je když uživatel nevyplní žádná historická data, ale přidá pouze pohyby v budoucnu. Zde by již bylo možné snažit se odhadnout budoucí vývoj, ale protože aplikace má provádět extrapolaci ze skutečně uskutečněných pohybů a ne z potenciálních, tak s této situací také nebude stanovena prognóza.

5.3.6 Datový případ – různé hodnoty bez periodicity

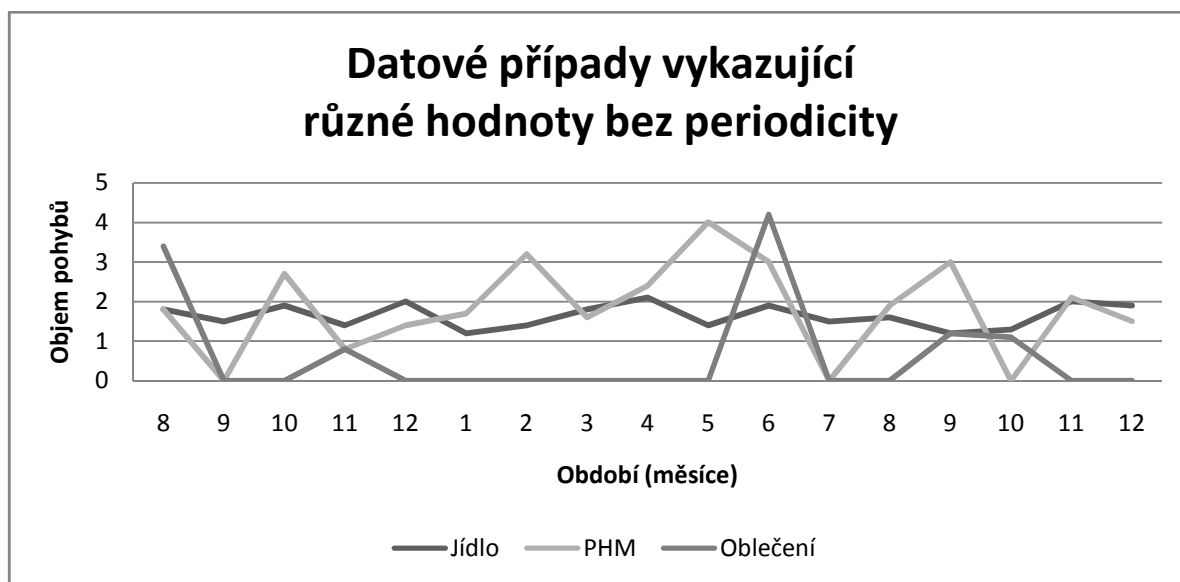
Časové řady budou zařazeny do této kategorie, pokud v datech nelze nalézt žádnou z výše popsaných charakteristik. Do tohoto datového případu bude spadat většina výdajů, které se neplatí paušálně, nesouvisí přímo s žádnou ze sezónních aktivit a neplatí se v přesně stanove-

ném intervalu. Typicky se může jednat o výdaje za potraviny, pohonné hmoty, výbavu domácnosti, oděvy a služby, jež se neplatí paušálně.

Tab. 6: Příklad dat nevykazujících žádné statisticky popsatelné chování.

Kategorie	Období (měsíce)															
	$t-16$	$t-15$	$t-14$	$t-13$	$t-12$	$t-11$	$t-10$	$t-9$	$t-8$	$t-7$	$t-6$	$t-5$	$t-4$	$t-3$	$t-2$	t
	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Jídlo	1,5	1,9	1,4	2	1,2	1,4	1,8	2,1	1,4	1,9	1,5	1,6	1,2	1,3	2	1,9
PHM	0	2,7	0,8	1,4	1,7	3,2	1,6	2,4	4	3	0	1,9	3	0	2,1	1,5
Příjem	0	0	75	8	14	0	0	0	54	0	2	0	31	1	4	0
Oblečení	0	0	0,8	0	0	0	0	0	0	4,2	0	0	1,2	1,1	0	0

Nepříjemnost při predikci tohoto typu časové řady je to, že se nechová žádným lehce popsatelným způsobem. Ve skutečnosti i nějaký klíč pro rozhodování uživatele být může, ale aplikace ho nedokáže popsat kvůli neznalosti principu, na jakém závislost funguje. Proto pro budoucí období nelze nabídnout měsíčně přesná data. To ovšem neznamená, že z takovýchto informací nelze uživateli sdělit užitečné informace pro budoucí plán. Přestože nemáme způsob jakým uživateli předpovědět, kdy si bude kupovat nové kalhoty, můžeme mu sdělit kolik v průměru utratí měsíčně za oblečení. Uživatel poté při sestavování plánu může počítat s částkou, jež by si měl na výdaje spojené s touto kategorií měsíčně spořit.



Obr. 7: Příklad dat nevykazujících silnou sezónnost.

Než bude časové řada prohlášena jako tento datový případ, měla by aplikace vyčerpat všechny ostatní možnosti, které dokážou nabídnout i měsíční rozlišení predikce. Další nevý-

hodou je, vzhledem k tomu že tato řada je definována jako nahodilá řada čísel, nelze definovat jasný detekční algoritmus. Proto je možné, že k tomuto případu budou přiděleny i kategorie silně nepravidelných výdajů, jež je lepší vůbec nepredikovat. Příkladem za všechny je plat z nezávislé činnosti. V tabulce č. 6 je příklad osoby vydělávající peníze nárazově. Člověk s tímto typem příjmu by se neměl spoléhat na průměr příjmu z minulosti, ale měl by si predikci pro tuto kategorii vypnout a své příjmy si do aplikace doplňovat podle skutečně získaných zakázek, či až po vystavení faktur.

Pro tento datový případ je velice důležité správné očištění. Pokud by v řadě zůstaly jednorázové pohyby, které spotřebitel provádí pouze výjimečně, byla by pak prognóza silně nepřesná. Proto by měl již uživatel takovéto pohyby vyřazovat do jiné kategorie pro níž prognóza prováděna nebude. Jako ochranu před pohyby, jež uživatel sám neodstraní, provede aplikace základní očištění. Ta bude provedena nad řadou, jež obsahuje pohyby alespoň v polovině období a nahradí všechny, jež přesahují více než trojnásobně průměr měsíců, v nichž se nachází pohyby. Tyto hodnoty jsou nahrazeny měsíčním průměrem z daného roku.

Po očištění můžeme uživateli sdělit částku, s níž by měl pro tuto kategorii měsíčně počítat. Musí ovšem být upozorněn, že se jedná pouze o aritmetický průměr a že se objem pohybů může meziměsíčně výrazně lišit.

5.4 Korelace dat

Při úvahách o možnosti prognózy peněžních toků byla stanovena i hypotéza, že se některé kategorie mohou navzájem ovlivňovat. Jako příklad byla situace, kdy jedinec při zvýšení platu začne více utrácet v restauraci. Bylo očekáváno, že také dvojice výdajů mohou vykazovat velkou míru závislosti. Pokud utrácíme za jednu oblast výdajovou kategorii, měl by se projevit úbytek jinde, protože objem finančních prostředků je limitovaný. Aplikace měla provést regresní analýzu mezi jednotlivými kategoriemi a poté při zjištění silné korelace upravit prognózu o tento vliv.

Mezi jednotlivými kategoriemi každého jedince byla prováděna korelační analýza. Jejím cílem bylo zjistit, zda se v těchto datech vyskytují případy, kdy jednotlivé výdajové a příjmové oblasti vykazují závislost. U dat byla zkoumána lineární i kvadratická korelace. Míra korelace byla měřena Pearsonovým korelačním koeficientem, který měl odhalit případnou závislost vystiženou lineární regresní funkcí a dále také indexem korelace, který vystihuje

závislost kvadratickou regresní funkcí. Na základě odhadu byl očekáván výskyt především lineární závislosti mezi některými příjmy a výdaji.

Po provedení analýzy bylo ovšem zřejmé že dvojice, vykazující silnou závislost, se běžně v osobních rozpočtech nevyskytují. V případech, kdy se zdálo, že data vykazují alespoň malou míru závislosti, se ukázalo, že jde pouze o náhodu. Pokusy promítnout jakkoliv korelaci do prognózy nepřinesly žádné zpřesnění predikce.

Příčin pro to, že nebyla nad daty naměřena žádná závislost, může být hned několik. Člověk může přírůstek příjmů utratit různými způsoby a při každém zvýšení či snížení příjmů se rozhodne jinak. Mezi dvojicemi výdajů také nebyla pozorována silná závislost. Velké množství kombinací výdajů na každý měsíc znemožnilo pozorování jasného vztahu.

6 Vlastní aplikace

6.1 Specifikace požadavků

Funkcí systému je nabídnout uživateli přehlednou formou prognózu vývoje jeho financí. Pro zajištění potřebných dat bude uživatel vkládat do systému jednotlivé pohyby, ovlivňující jeho osobní bilanci. Tyto pohyby bude uživatel dělit do jednotlivých datových kategorií, podle svého úsudku. Těchto kategorií může být libovolný počet a mohou obsahovat libovolné položky. Výstup pro uživatele bude formou předpokládaných příjmových a výdajových částek, které budou jasně zobrazeny pro každý měsíc a datovou kategorii. Vzhledem k tomu že ne všechny typy pohybů je prognóza žádoucí a účinná, musí ji mít uživatel možnost pro jednotlivé kategorie zapnout či vypnout. Aplikace musí být dostupná v prostředí internetu a nesmí vyžadovat žádnou instalaci na pracovní stanici.

6.1.1 Účel a využití aplikace

Účelem této aplikace je poskytnout uživateli jednoduchý a užitečný nástroj pro analýzu osobní cash flow bez využití externího finančního poradce. Aplikace by měla být použitelná kterýmkoliv uživatelem bez rozdílu výše obrátu či dosaženého vzdělání v oboru financí. Výstupem aplikace by měla být lehce uchopitelný plán vývoje cash flow.

6.1.2 Základní popis aplikace

Aplikace by měla sbírat a uchovávat data o peněžních pohybech jedince v průběhu času. Po sběru by měla data vyhodnotit a uživateli přehledně předložit předpokládaný vývoj jeho hoto-
vostních toků v horizontu jednoho roku za pomoci dostupných metod extrapolací analýzy. Tato aplikace by měla kvůli dostupnosti využívat prostředí internetu.

6.1.3 User stories

Abychom lépe porozuměli tomu, co má aplikace uživatelům nabídnout, nadefinujeme si několik „user stories“. Tyto krátké příběhy, často používané při agilních metodikách vývoje, popisují, co uživatelé tímto systémem chtějí dosáhnout.

- Sociálně slabší uživatel chce zjistit, kolik může utratit za jednotlivé položky tak, aby v následujících měsících vyšel s penězi.

- Uživatel, který má dle svého názoru dostatek peněz, chce zjistit, zda může investovat přebytek rozpočtu a kolik to je, při zachování dostatečné rezervy.
- Uživatel chce spořit maximální částku svého rozpočtu a hledá, v kterých výdajových položkách by mohl ušetřit.

Z těchto user stories vyplívají dva hlavní typy uživatelů a to jsou ti, kteří chtějí zajistit osobní rozpočet bez deficitu a poté ti, kteří nemají problém s likviditou, ale potřebují ucelené informace o svých příjmech a výdajích pro svoje další rozhodnutí.

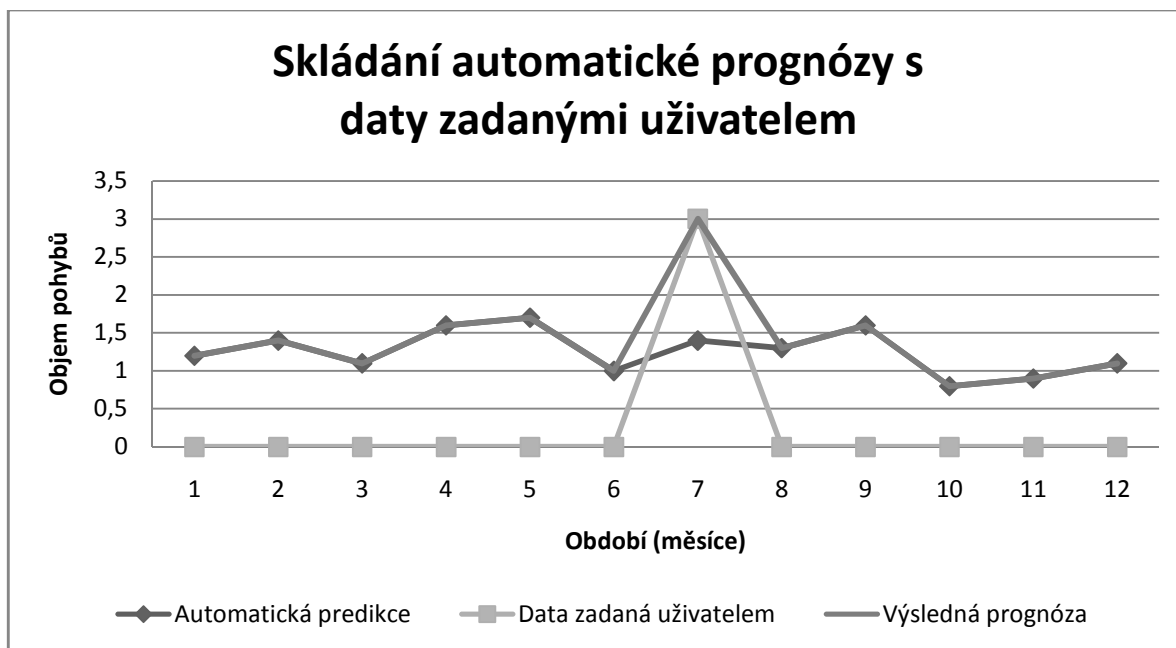
6.1.4 Uživatelé

Aplikace bude navržena především pro fyzické osoby, které požadují vypracování předpokládaného vývoje osobních financí. Z povahy aplikace bude ovšem možné, aby ji použila i firma analyzující svůj budoucí vývoj finančních prostředků, či si prohlédla vývoj prodeje jednotlivých výrobků ve vztahu k sezónním vlivům.

6.1.5 Vkládání dat k budoucím pohybům

Až do této fáze jsme uvažovali jak strojově předpovědět vývoj cash-flow v budoucím roce. Ať bude ovšem algoritmus jakkoliv dokonalý, nikdy nepodchytí nečekané výdaje a zvláštní situace, jež nejsou implementovány. Pro plánování je proto naprosto nezbytné, aby měl uživatel možnost ovlivňovat budoucí rozpočtový plán pomocí vstupů nejen z historie, ale i do budoucnosti. Uživatel proto musí mít způsob, jakým vkládat známe pohyby v budoucnu. Kromě toho, že bude moci zvýšit objem pohybů v daném měsíci, měl by mít možnost i nastavit nižší objem, pokud si je například jistý, že v daném měsíci nebudou v této kategorii žádné výdaje. Poté ovšem musí algoritmus rozprostřít předpokládanou částku do ostatních období.

Tato funkce do aplikace poté přináší otázku, jakou hodnotu zvolit při práci s budoucími daty. Toto je například důležité při korelaci s jinými kategoriemi.



Obr. 8: *Skládání prognóz.*

V grafu na obrázku 8 vidíme projekci budoucích výdajů na pohonné hmoty spočítanou podle osobních historických dat. Příští rok se ovšem uživatel rozhodl, že použije automobil k letošní dovolené, a proto mu červenci prudce vzrostou náklady na benzín. V jiných měsících si ale uživatel přeje nadále počítat s prognózou z minulých let. Naprosto přirozeně proto přidá tento zvýšený odhad na tento měsíc. Aplikace poté zahrne tyto data do budoucí výsledné prognózy.

Pro správné určení hodnoty pro budoucí měsíce musíme vždy uvažovat možnost, že data pro budoucí měsíc uživatel nevyplnil kompletní. Pokud je tedy částka strojové prognózy větší než objem pohybů zadaných uživatelem, měla by se nadále výsledná prognóza počítat z objemu spočítaného aplikací.

6.1.6 Prezence a vysvětlení výsledků uživateli

Aplikace je určena především uživatelům, kteří nemají zkušenosti s finančním plánováním. Právě proto je velice důležité správně vysvětlit výsledky prognózy uživateli. Ten musí pochopit charakter předpokladu pro následující období, aby měly data při plánování osobního rozpočtu smysl. Uživatel by měl být také upozorněn na to, kdy předpověď všeobecně nedosahuje dobrých výsledků (např. u kategorií s velice nepravidelnými výdaji), či na to že nepravidelné velké výdaje jako jsou nákupy nemovitostí, či zařízení domácnosti, predikce nemůže ze své podstaty odhalit vůbec. Uživatel by měl výsledky predikce vidět jako měsíční objemy pro

každou kategorii, měsíční součet za každý měsíc a poté také jako spojnicový graf vývoje volných peněžních prostředků, který uživateli pomůže v lepší orientaci ve vývoji svého rozpočtu.

6.1.7 Uživatelské rozhraní

Do styku s aplikací se budou dostávat i nezkušení uživatelé a je proto nutné aby aplikační rozhraní bylo co nejvíce intuitivní a přirozené. Uživatele není možno na práci se systémem nijak školit, a proto musí být jednoduchost aplikace na prvním místě.

6.1.8 Bezpečnostní požadavky

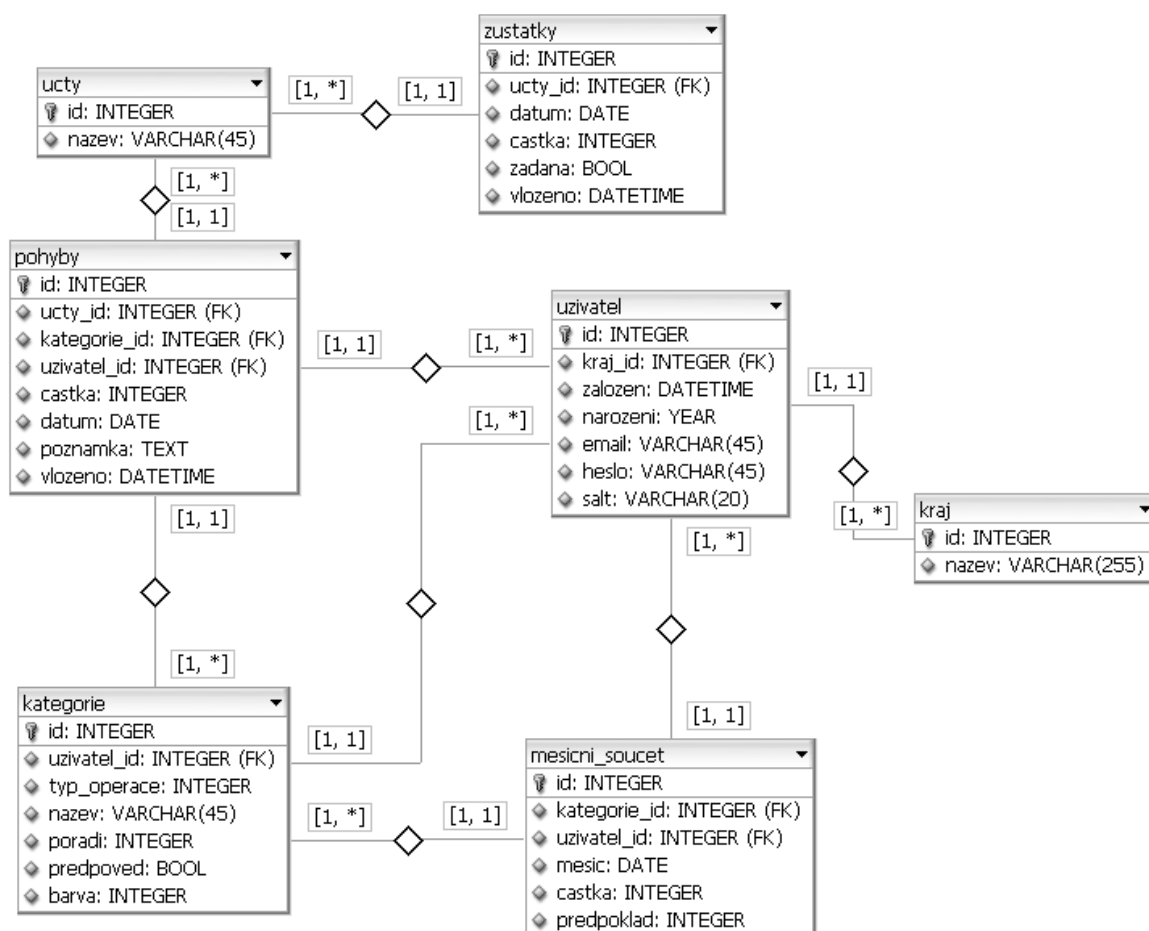
Systém bude uchovávat údaje o finančních transakcích uživatelů a je proto třeba aplikovat co nejvíce opatření pro zabránění úniku dat a jejich případnému zneužití neoprávněnou osobou. Aplikace by proto neměla od uživatelů požadovat žádné osobní údaje, které se dají použít pro přiřazení konkrétní osoby k danému uživatelskému účtu.

6.1.9 Architektura a technické řešení aplikace

S požadavků na systém vzešlo řešení, kdy se u uživatelů nesmí vyžadovat instalace žádné další aplikace, a proto je aplikace navržena jako webový systém. Pro vývoj aplikace je dostupný webový server, který nabízí podporu jazyka PHP a disponuje databázovým serverem MySQL. Proto i aplikace bude naprogramována a odladěna v jazyce PHP 5 a na databázovém serveru MySQL 5.1.

6.2 Databázový návrh

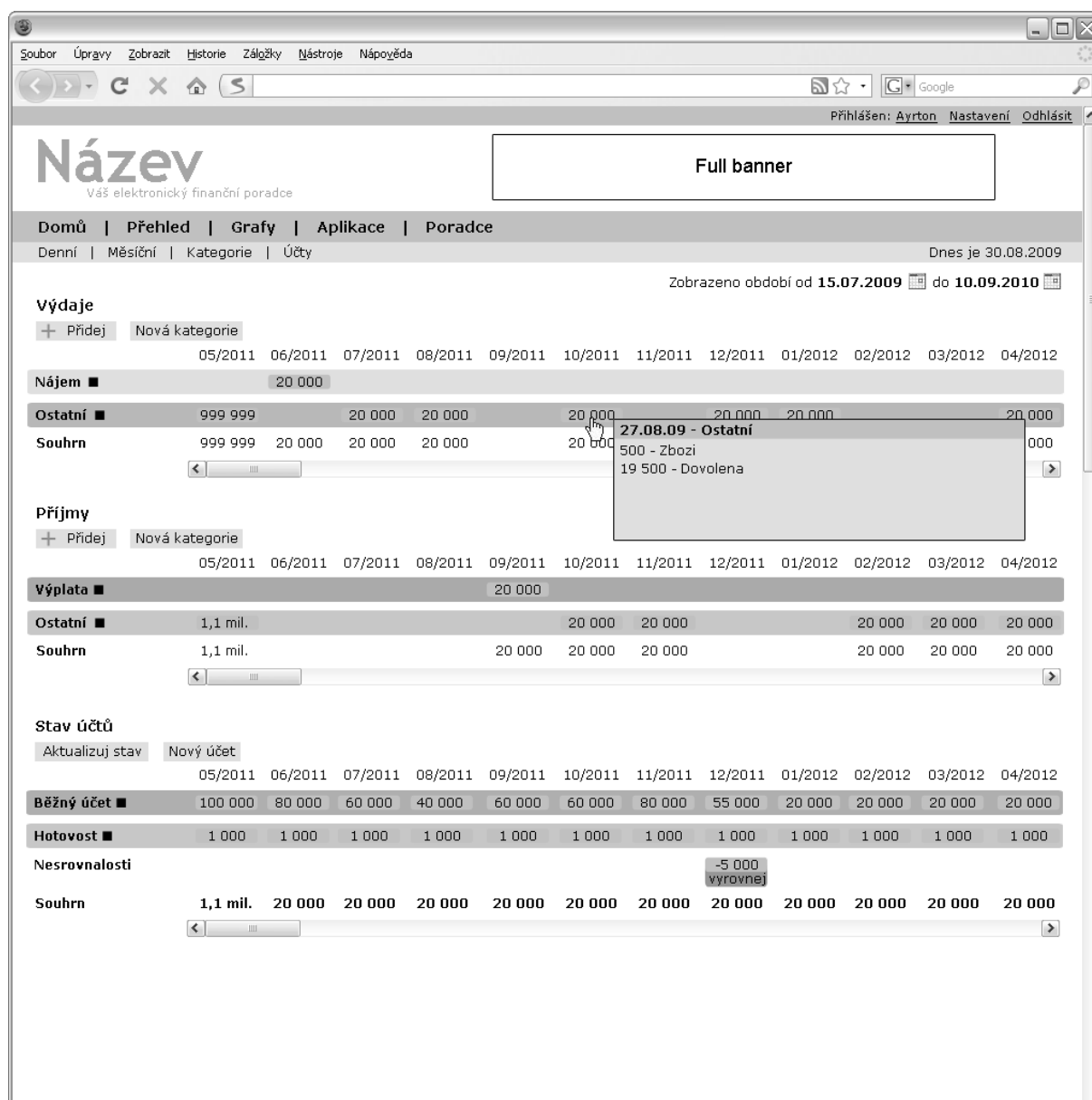
Schéma databáze aplikace bylo navrženo s důrazem na jednoduchost a rozšiřitelnost. Kromě tabulek udržujících data o uživatelích a jejich peněžních pohybech, je zde navržena také tabulka udržující výsledky automatické prognózy. Díky tomu není potřeba provádět výpočet při každém načtení stránky a predikce se tak může provádět jen při přidání, editaci či mazání dat.



Obr. 9: Schéma databáze

6.3 Podoba aplikace

Pro užitečný webový systém pro plánování osobních financí, je třeba zde navrženou aplikaci zasadit do širší aplikace, jež bude uživatelům poskytovat i jiné nástroje než pouze krátkodobou predikci cash-flow. Proto zde není navržena finální grafická podoba celého systému, ale pouze tzv. wireframe, jež je zobrazením obsahu a rozmístění ovládacích prvků aplikace.



Obr. 10: Wireframe aplikace

6.4 Implementace algoritmu predikce

Algoritmus predikce je v aplikaci implementován podle všech výše zjištěných poznatků. Postup je následující. Nejprve aplikace zhodnotí, zda za posledních 12 měsíců byly zaznamenány jakékoliv pohyby u dané kategorie. Pokud ne, aplikace neprovádí žádnou projekci. Dá se předpokládat, že kategorie je nepoužívaná, či naopak čerstvě založena a neobsahuje tedy ještě žádná data. V opačném případě aplikace začne zkoumat míru sezónnosti. Podle délky dat je to buď sezónnost roční, pololetní resp. čtvrtletní. Speciálním případem je velice krátká řada, kdy aplikace disponuje daty pouze za aktuální měsíc, u té aplikace pouze rozkopíruje tento měsíc-

ní údaj na příští měsíce. Pokud je ukazatel sezónnosti nad stanovenou hranicí, uvedenou výše v textu, je provedena predikce na základě metody extrapolace se zohledněním sezónnosti. V opačném případě se zjistí, zda kategorie vykazuje neměnnou hladinu či schodovité chování, a pokud ano, pak aplikace rozkopíruje data z minulého měsíce do každého měsíce pro příští rok. Pokud data nevykazují žádnou z popsaných charakteristik, je provedena extrapolace pouze na základě aritmetického průměru.

6.5 Test úspěšnosti predikce

Po sestavení metody extrapolace jednotlivých kategorií, byla změřena přesnost s jakou je schopna predikovat jednotlivé kategorie. Měření probíhalo tak, že systému byly nabídnuty časové řady, ze kterých bylo vyjmuto posledních 12 měsíců. Tyto vyňatá data za posledních dvanáct měsíců byla poté srovnána s extrapolovanými hodnotami z predikce. Odchylka od reality byla měřena vlastním ukazatelem přesnosti, jehož metodiku výpočtu jsme si definovali výše. Kategorie, jež byly použity v měření přesnosti, musely obsahovat data minimálně za 16 měsíců, aby byl vyloučen vliv predikcí velice krátkých řad. Navíc byly z měření vyloučeny řady, u nichž respondenti označili, že u nich není predikce vhodná.

Na míru přesnosti mělo vliv několik faktorů. Nejsilnější vliv na přesnost predikce měla délka historických dat. S přibývajícím množstvím dat se přesnost pochopitelně zvyšovala. Záleželo také na tom, kolik měsíců z roku obsahovalo peněžní pohyby. Čím méně bylo v roce vyplněných měsíců, tím se přesnost zhoršovala, obzvláště pokud takováto řada nevykazovala silnou sezónnost. U takovýchto řad by uživatel měl stejně predikci vypnout a zadávat data manuálně, proto by tento specifický případ neměl predikci nijak silně ovlivňovat.

Po analýze byla naměřena 88% míra přesnosti. Vzhledem k tomu, že všechny časové řady byly vyjádřeny ve stejné měně, můžeme si i vyjádřit, že u testovaného datového vzorku se rovnala průměrná jedenáctiprocentní nepřesnost 2 061 špatně alokovaným korunám měsíčně.

6.6 Porovnání predikcí

Po analýze úspěšnosti navržené metody predikce byla míra úspěšnosti porovnána s odhadem uživatelů, který uvedli v prvním kole průzkumu. Porovnání probíhalo na datech za 6 měsíců, které byly získány ve druhém kole ankety. Zatímco popsaná metoda za toto období dosáhla přesnosti 87%, uživatelé dosáhli pouze 79% přesnost. Je zde třeba poznamenat, že všichni respondenti si již nějakým způsobem peněžní deník vedou a proto mají relativně dobrý pře-

hled o výši svých příjmů a výdajů. Jakou přesnost dosahují lidé, jež si nevedou žádnou formou záznamy o svých financích, zkoumáno nebylo. Je ale zřejmé, že přesnost jejich odhadu by byla oproti respondentům ankety výrazně horší.

Důvodem proč uživatelé dosahují horších výsledků je především to, že správně neodhadnou průměrnou částku, kterou vydávají za konkrétní kategorii. Tento špatný odhad je poté hlavním důvodem, proč vykazují menší přesnost než navržený algoritmus.

6.7 Finanční aspekty provozu aplikace

Aplikace skýtá velký potenciál pro komerční využití. Ovšem i pokud budeme uvažovat, že si uživatelé aplikaci oblíbí a budou ji opravdu využívat, pak vyvstává otázka, z čeho přesně financovat vytvoření a provoz aplikace. Před spuštěním aplikace do ostrého provozu by musela aplikace obsahovat mnoho funkcí a být použitelná i na mobilních platformách. Zároveň by musel být vystavěn systém nápovědy a podpory, pokud by měl systém obstát v konkurenci existujících nástrojů pro finanční plánování.

Problém financování můžeme proto rozdělit na dvě různé oblasti.

- Jak zajistit finanční prostředky na vystavění aplikace do stádia, kdy stojí v konkurenci existujících aplikací.
- Jak po dosažení dostatečné návštěvnosti zajistit finance na provoz a údržbu systému.

Problém prvotní investice je v dnešní době vysokého zájmu o internetové projekty celkem dobře řešitelný. Získání investora, jež podpoří vývoj aplikace finančně i zkušenostmi, by nemělo být nepřekonatelným problémem. Mnoho takovýchto „obchodních andělů“ (angl. business angels), jak se těmto investorům vyhledávajícím atraktivní IT firmy v počátcích říká, působí v zahraničí a pár již také v tuzemsku. Prodej podílu na firmě, vyvíjející takovýto systém, by měl proto být dostatečný pro rozvoj aplikace.

Pokud se prokáže, že aplikace stojí v konkurenci ostatních nástrojů osobního finančního plánování, je poté důležité vyřešit otázku, jak bude firma provozující tento systém financovat jeho údržbu, další vývoj a získávat případný zisk. Pokud budeme uvažovat dostatečnou návštěvnost aplikace, nabízí se několik možností výdělku pro provozovatele takového systému.

- **Zpoplatnění uživatelského přístupu či některých funkcí.** První nápad při zkoumání možností příjmu z internetových služeb je zpoplatnění registrace uživatelů. Takovýto přístup je velice jednoduchý, ale pro systém představuje velký risk. Mnoho uživatelů očekává od služeb na internetu, že jsou zadarmo. Proto se aplikaci může stát, že po zavedení tohoto systému ztratí mnoho uživatelů a získávání nových bude velice složité.
- **Reklama.** Další běžně používaný nástroj pro zpeněžení internetové služby je jakákoliv forma reklamy. Aplikace je jasně zacílena na uživatele, kteří mají zájem o správu svých financí a přehled jejich stavu, čímž se stává atraktivní pro celou řadu inzerentů. Na druhou stranu ale nejnovější výzkumy zpochybňují účinnost reklamy a příjmy z ní se proto neustále snižují.[7]
- **Prodej dat.** Dalším, méně běžným způsobem výdělku na internetu je prodej agregovaných dat či ukazatelů z nich. Aplikace obsahuje přesné informace o výdajích a příjmech uživatelů a dozvídá se je v reálném čase. Takováto data mohou mít pro nejrůznější firmy, výzkumníky a státní instituce obrovskou hodnotu. Při hledání dat pro testování predikčního algoritmu nabyly nalezeny ani žádná historická data natož pak taková, která jsou k dispozici v reálném čase. Je si ovšem potřeba uvědomit, že jakékoliv nakládání s daty jsou uživatelé aplikace velice citliví a bylo by proto třeba provádět veškeré obchodní aktivity velice citlivě s obrovským důrazem na soukromí uživatelů. Data by proto musela být zbavena všech položek, jež by mohly poukázat na identitu vlastníka. Pravděpodobně by bylo možné nabízet pouze jakési výběry z dat, či ukazatele spočítané z nich.
- **Zprostředkování prodeje finančních produktů.** Například největší aplikace na trhu Mint si vydělává právě tímto způsobem.[8] Tato služba doporučuje produkty bank a při úspěšném přechodu klienta poté získává provizi. Tento systém umožňuje velký potenciál pro výdělek, ovšem závisí na dobré znalosti národního prostředí a existenci smluv s místními poskytovateli finančních produktů. Nelze jej jednoduše provádět na mezinárodní úrovni.

Aplikace by pro svůj běh měla používat kombinaci uvedených přístupů. Celý algoritmus predikce je také možno zpeněžit tím, že by se modul pro automatizovanou predikci prodal jedné ze zavedených firem zabývajících se vývojem aplikací pro osobní finance. V podobném

duchu se pak také nese přístup, kdy se firma nesnaží mít žádný provozní zisk, ale místo toho se soustředí pouze na získávání uživatelů za peníze z prodeje akcií. Investoři v tomto případě spoléhají na to, že firmu bude v pozdější fázi chtít zakoupit ještě silnější investor. Tento přístup v posledních letech získává na popularitě, závisí ovšem na velkém zájmu investorů.

7 Diskuse

V této práci je řešen pouze problém elektronického plánování cash flow, ale systém, který slouží ke správě osobních financí, musí obsahovat mnohem více. Systém musí uživateli pomoci s finančními rozhodnutími pro celý život. Proto je třeba, aby komplexní aplikace pro finanční správu, obsahovala také systém dlouhodobého plánování. Tato problematika je velice složitá. Do hry tu vstupuje proměnlivá hodnota peněz v čase, různá likvidita aktiv a také rizika, s kterými se lidé setkávají. Aplikace by měla uživatelům pomoci při jejich investičních rozhodnutích tak, aby výnosy byli k dispozici ve správný moment. Zároveň by také měl systém pomoci s finanční přípravou na stáří.

V práci bylo také uvažováno o zapojení neuronových sítí. Použití neuronových sítí v aplikaci, jež uživatelům neustále nabízí aktuální predikci cash flow, je ovšem značně problematické. Díky rozdílnému chování spotřebitelů by bylo vhodné, aby probíhalo učení neuronové sítě vždy s daty konkrétního uživatele. Takovýto přístup by byl nepoužitelný v situacích, kdy uživatel nezadal do systému velké množství dat a je nepravděpodobné, že by se systémem pracoval až do této chvíle bez jakékoliv predikce. V definici požadavků bylo navíc uvedeno, že aplikace musí provádět predikci nezávisle na měně, což je také zásadní překážkou při použití neuronových sítí. Síť naučená na jedné měně by selhávala při predikci s měnou jinou. Stanovení rozmezí by bylo také značně problematické. Jednou variantou by byl jakýsi hybridní přístup, při kterém by se používala metoda definovaná v této práci a teprve po dosažení velkého množství dat by se použila predikce se zapojením neuronové sítě.

8 Závěr

Na počátku této práce stála hypotéza, že je možné, aby uživatel prostým vedením finančního deníku a zapojením automatizovaného systému predikce peněžních toků výrazně zpřesnil svůj finanční plán. Tato byla potvrzena a zároveň byla navržena metoda, jež ukazuje jeden ze způsobů jak toho dosáhnout. Zároveň byla navržena aplikace, jež by tuto metodu využila za použití definovaných zásad.

Při konzultacích u sběru ankety bylo zjištěno, že i uživatelé, kteří si nějakým způsobem vedou finanční deník, nemají přesné informace, kolik průměrně utrací za jednotlivé položky v rozpočtu, a proto jsou jejich odhady značně nepřesné. Právě v tomto může automatizovaná predikce uživatelům pomoci a ti při jejím použití v kombinaci se svým úsudkem mohou dosáhnout vyrovnaného plánu osobních financí.

Samostatná aplikace by při ostrém nasazení neměla šanci obstát. Tato by měla být implementována jako součást systému poskytujícího více nástrojů finančního plánování, jako jsou aplikace dlouhodobého (celoživotního) plánování, rozpočtování apod. Je proto nutné najít finanční prostředky pro vývoj rozsáhlejšího systému, nebo nabídnout tuto aplikaci již existujícím službám, z nichž žádná takovouto funkčnost nenabízí.

9 Použitá literatura

1. ARTL, Josef; ARTLOVÁ, Markéta. *Ekonomické časové řady*. 1. vydání. Praha : Professional Publishing, 2009. 290 s. ISBN 978-80-85946-85-6.
2. The need for financial planning. *Personal finance* [online] 2011-11-02, 11, [cit. 2011-10-09]. Dostupný z WWW: <<http://www.rediff.com/money/2007/nov/02perfin1.htm>>.
3. Consumer Expenditure Survey - Composition of consumer unit. In *Bureau of labor and statistics*. – United States Department of Labor 2010
4. ARSHAM, Hossein; SHAO, Jr. Seasonal and Cyclic Forecasting for the Small Firm. *American Journal of Small Business*. 1985-04-01, 9, 4, s. 46-57. Převzato z kolekce Business Source Complete. ISSN 03639428.
5. HINDLS, Richard; HRONOVÁ, Stanislava; NOVÁK, Ilja. *Metody statistické analýzy pro ekonomy*. 2. vydání. Praha : Management Press, 2000. 259 s. ISBN 80-7261-013-9.
6. SEDLÁČEK, Jaroslav. *Cash Flow*. První vydání. Brno : Computer press, 2003. 189 s. ISBN 80-7226-875-9.
7. CHTOUROU, Mohamed Saber, et al. Is banner blindness genuine? Eye tracking internet text advertising. *Applied Cognitive Psychology* [online]. 2011-09-01, 25, 5, [cit. 2011-11-23]. s. 708-716. Převzato z kolekce Academic Search Complete. Dostupný z WWW: <<http://search.ebscohost.com/login.aspx?direct=true&db=a9h&an=65551937>>. ISSN 08884080, eclanek 10.1002/acp.1742.
8. *Mint* [online]. 2011 [cit. 2011-11-23]. Financing Mint. Dostupné z WWW: <<https://www.mint.com/how-it-works/free/>>.
9. HUSEBY, S. - *A Security Wake-Up Call for Web Programers* 1. vyd. New York : John Wiley sons Inc, 2003. 265 s., ISBN 978-0-470-85744-1
10. WESSLING, Harry - *Aktivní vztah k zákazníkům pomocí CRM* 1. vydání. Praha: Grada publishing, 2003, ISBN 80-247-0569-9.
11. FIGHT, A. *Cash Flow Forecasting*. Oxford: Butterworth-Heinemann, Elsevier Science, 2005. 256 s. ISBN 0-7506-6136-4.
12. MAŘÍK, V. -- ŠTĚPÁNKOVÁ, O. -- LAŽANSKÝ, J. a kol. *Umělá inteligence* 1. 1. vyd. Praha: Academia, 1993. 264 s. ISBN 80-200-0496-3.
13. PARNAS, David; CLEMENTS, Paul. A Rational Design Process: How and Why to Fake It. *Transactions on software engineering*. 1986, 2., s. 251 - 256. Dostupný také z WWW: <<http://www.idemployee.id.tue.nl/g.w.m.rauterberg/presentations/parnas-clements-1986.pdf>>.