**VERJETNOST PREKOMERNEGA PRILEGANJA**

**Motivacija:** Ena od glavnih težav, s katero se srečujemo pri testiranju kvantitativnih trgovalnih strategij, je nevarnost prekomernega prileganja (overfitting) modela na izbrani učni množici. Izvajanje strategije na testnih množicah predstavlja le delno rešitev problema.

**Moja naloga**: za dani nabor trgovalnih strategij izračunaj verjetnost prekomernega prileganja z novo metodo kombinatoričnega simetričnega navzkrižnega preverjanja.

**Opis dela:**

**1. podatki:**

Najprej sem pridobila podatke za delnice, ki so trenutno v indeksu S&P 500. Za vseh 500 delnic sem vzela zadnjo ceno (Close vrednost). Podatke sem pridobila iz spletne strani yahoo finance s pomočjo funkcije v R. Podatke sem vzela za čas od 1. januarja 2000 do 1. novembra 2013 (približno 13 let).

Indeks SP500: <http://en.wikipedia.org/wiki/List_of_S%26P_500_companies>, <http://en.wikipedia.org/wiki/S%26P_500>

Standard & Poor's 500 (oz. S&P 500) je borzni indeks, ki temelji na tržni kapitalizaciji 500 velikih podjetji, katerih navadne delnice kotirajo na NYSE (Newyorška borza) ali NASDAQ (National Association of Securities Dealers Automated Quotations).

Tiste delnice, ki ne obstajajo dovolj dolgo, sem izločila. Izločila sem tudi tiste delnice, ki imajo več kot tri manjkajoče vrednosti skupaj. Ostale manjkajoče vrednosti sem nadomestila z vrednostjo zadnje cene prejšnjega trgovalnega dne.

**2. strategije:**

Nato sem sestavila nekaj enostavnih trgovalnih strategij:

Nekatere na podlagi matematičnih indikatorjev (<http://www.markets.com/sl/education/technical-analysis/math-indicators.html>)

SMA (simple moving average): Preprosto drseče povprečje je aritmetična sredina gibanja cene. Izračuna se tako, da seštejemo vse vrednosti v danem obdobju, potem pa vse skupaj delimo z željeno dolžino obdobja. Pri izračunu vedno upoštevamo najkasnejša obdobja. Pogosto se uporablja kombinacija SMA 5, 10, 20, 50, 100 in 200, pri čemer številka predstavlja dolžino iskanega obdobja. (<http://pecuniati.wordpress.com/2008/01/14/preprosto-drsece-povprecje-sma/>)

Drseča povprečja se prikažejo na istem grafu kot gibanje cen. Nakupni znak nam pokaže takrat, ko se končna cena povzpne nad krivuljo drsečega povprečja, prodajni pa, ko pade pod njo. Daljša drseča povprečja so pomembna za določanje dolgoročnih trendov.   
Krajša povprečja pa so primernejša pri nihanju cen na določenem intervalu kakor pri trendnem gibanju cen. Vendar so pri obratu trenda zopet uporabnejša krajša povprečja, saj daljša zaznajo obrat trenda z večjo zamudo, ker so manj občutljiva in bolj oddaljena od dejanskega gibanja cene.  
Uporaba dolžine povprečne drseče sredine je subjektivna in je odvisna od vsakega posameznika. Za kratko obdobje se najpogosteje uporablja 5, 10 in 20 dnevne drseče sredine, med tem ko za daljše obdobje analitiki najbolj pogosto uporabljajo 50, 100 in 200 dnevno drsečo sredino.   
Signale za nakup in prodajo nam poda tudi križanje drsečih sredin. V primeru da krajša povprečna sredina (npr. 50 dnevna) prebije daljšo povprečno sredino (npr. 100 dnevno) od spodaj navzgor je to signal za nakup in obratno signal za prodajo. Za potrditev signala za nakup ali prodajo navadno nista dovolj dve drseči sredini, večina trgovalcev uporablja pet (običajno 5, 10, 20, 100 in 200 dnevne) pri čemer pomembnost sekanj z daljšim obdobjem narašča. Pogosto se zgodi, daje ta signal prepozen zato ga je priporočljivo uporabljati s kakšnim dopolnilnimi indikatorji, kot sta med drugim tudi RSI, MACD. (<http://www.dnevnik.si/poslovni/novice/1042237812>)

RSI (Relative Strength Indeks): Indikator indeks relativne moči oziroma Relative Strength Index (RSI), ki ga je razvil J. Welles Wilder je eden izmed najbolj uporabljenih in uporabnih vodilnih indikatorjev. Vodilni pomeni, da nam da signal pred samo spremembo trenda gibanja cene delnice. Najbolje se uporablja v kombinaciji s črtami trenda, vzorci ter črtami podpore in odpora.

RSI primerja relativno moč vsote dobičkov v dneh preučevanega obdobja, ko je cena naraščala (to je v dneh, ko je bila končna cena dneva višja od končne cene prejšnjega dneva) v primerjavi z vsoto izgub v dneh proučevanega obdobja, ko se je cena znižala (to je v dneh, ko je bila končna cena dneva višja od končne cene prejšnjega dneva).

**Primerja velikosti rasti z velikostmi padcev posameznih vrednostnih papirjev**

Bistvo tega indikatorja je, da primerja velikosti rasti z velikostmi padcev posameznih vrednostnih papirjev. Indeks se giblje med vrednostima 0 in 100. Če vrednost indeksa pade pod 30, je to znak za podcenjenost vrednostnega papirja in bo verjetno sledil porast cene, če pa zraste čez 70, je to znak za precenjenost vrednostnega papirja in sledil naj bi padec cene. Nekateri analitiki prilagajajo meji in dolžino obdobja vsakemu trgu posebej. Pri daljših obdobjih upoštevajo ožje meje.

Kazalnik RSI se lahko izračunava za različno dolga obdobja. Wilderjev izvirni predlog je bil 14 dni, danes pa večina tehničnih analitikov uporablja občutljivejše kazalnike (5-dnevne, 7-dnevne ali 11-dnevne RSI). Čim daljše je obdobje, tem stabilnejši je kazalnik in manj je signalov, tudi napačnih. (<http://www.finance.si/186616>)

**Indeks relativne moči (RSI)**

RSI je oscilator, ki meri trenutno **moč tečaja** glede na tečajno zgodovino. Indikatorja RSI ne smemo zamenjati za relativno moč, ki meri moč instrumenta glede na moč trga.

RSI ima vrednosti med 0 in 100. Formula za izračun je naslednja: 100-(100/(1-A)), kjer je A povprečna vrednost pozitivnih period v opazovanem obdobju deljeno s povprečno vrednostjo negativnih period. Pozitivne in negativne periode so izračunane kot absolutna sprememba tečaja od zaključnega do zaključnega tečaja. ( <http://www.finport.si/Dokument.aspx?ID=65>)

Bollinger: **Bollinger Bands** is a [technical analysis](http://en.wikipedia.org/wiki/Technical_analysis" \o "Technical analysis) tool invented by [John Bollinger](http://en.wikipedia.org/wiki/John_Bollinger) in the 1980s, and a term [trademarked](http://en.wikipedia.org/wiki/Trademark" \o "Trademark)by him in 2011.[[1]](http://en.wikipedia.org/wiki/Bollinger_Bands#cite_note-1) Having evolved from the concept of trading bands, Bollinger Bands and the related indicators %*b* and *bandwidth* can be used to measure the "highness" or "lowness" of the price relative to previous trades. Bollinger Bands are a [volatility](http://en.wikipedia.org/wiki/Volatility_(finance)" \o "Volatility (finance)) indicator similar to the [Keltner channel](http://en.wikipedia.org/wiki/Keltner_channel" \o "Keltner channel).

Bollinger Bands consist of:

* an [*N*](http://en.wikipedia.org/wiki/Variable_(mathematics)#Naming_conventions)-period [moving average](http://en.wikipedia.org/wiki/Moving_average_(finance)" \o "Moving average (finance)) (MA)
* an upper band at *K* times an *N*-period [standard deviation](http://en.wikipedia.org/wiki/Standard_deviation) above the moving average (MA + *Kσ*)
* a lower band at *K* times an *N*-period [standard deviation](http://en.wikipedia.org/wiki/Standard_deviation) below the moving average (MA − *Kσ*)

Typical values for *N* and *K* are 20 and 2, respectively. The default choice for the average is a simple [moving average](http://en.wikipedia.org/wiki/Moving_average), but other types of averages can be employed as needed. [Exponential moving averages](http://en.wikipedia.org/wiki/Moving_average_(finance)#Exponential_moving_average) are a common second choice.[[note 1]](http://en.wikipedia.org/wiki/Bollinger_Bands#cite_note-3) Usually the same period is used for both the middle band and the calculation of standard deviation.[[note 2]](http://en.wikipedia.org/wiki/Bollinger_Bands#cite_note-4)(<http://en.wikipedia.org/wiki/Bollinger_Bands>)

ollingerjeve trakove je razvil John Bollinger na začetku osemdesetih let dvajsetega stoletja. Uporabljajo se za identifikacijo cenovnih vrhov in dna. Bollinger je razumel potrebo dinamičnih, prilagodljivih se trgovskih trakov, katerih delitev se razlikuje v odvisnosti od razpršenosti cen. V času visoke razpršenosti se Bollingerjevi trakovi širijo in kot takšni dajejo več prostora. V času nizke razpršenosti pa se ožijo, da ohranijo cene v svojih mejah.

* **Izračun**

Bollingerjevi trakovi so sestavljeni iz kompleta treh krivulj v odnosu do cen.

Srednja krivulja predstavlja srednjeročni trend. Temu namenu je navadno namenjeno dvajsetdnevno preprosto drseče  povprečje (SMA).

Zgornja meja traku se vodi enako kot srednja krivulja, samo, da je premaknjena navzgor za dva standardna odklona. Standardni odklon je vzorec, ki meri razpršenost in prikazuje, kako se cene razlikujejo od dejanske vrednosti.

Spodnja meja se ponovno vodi enako kot srednja krivulja, ampak je premaknjena za dva standardna odklona navzdol tako, da odgovarja razpršenosti trga.

Bollingerjevi trakovi ustvarjajo trak cenovnega razpona, določeno relativno merilo, ki meri kje je aktualna cena z vidika trakov.

**Spodnji Bollingerjev trak = SMA - 2 standardna odklona**

**Zgornji Bollingerjev trak = SMA + 2 standardna odklona**

**Srednji Bollingerjev trak = 20-dnevno preprosto drseče povprečje (SMA)**

* **Pomen**

Ko se širina traku zoži, obstaja verjetnost naraščanja hitrega uhajanja cen.

Če cene stalno napadajo zgornjo mejo Bollingerjevega traku, se orodje smatra kot prekupljeno in to lahko pomeni signal za prodajo.

Če pa se, obratno, cene stalno dotikajo spodnje meje traku, se smatra vrednostni papir za preprodan, kar je signal za nakup. (<http://www.markets.com/sl/education/technical-analysis/math-indicators.html>)

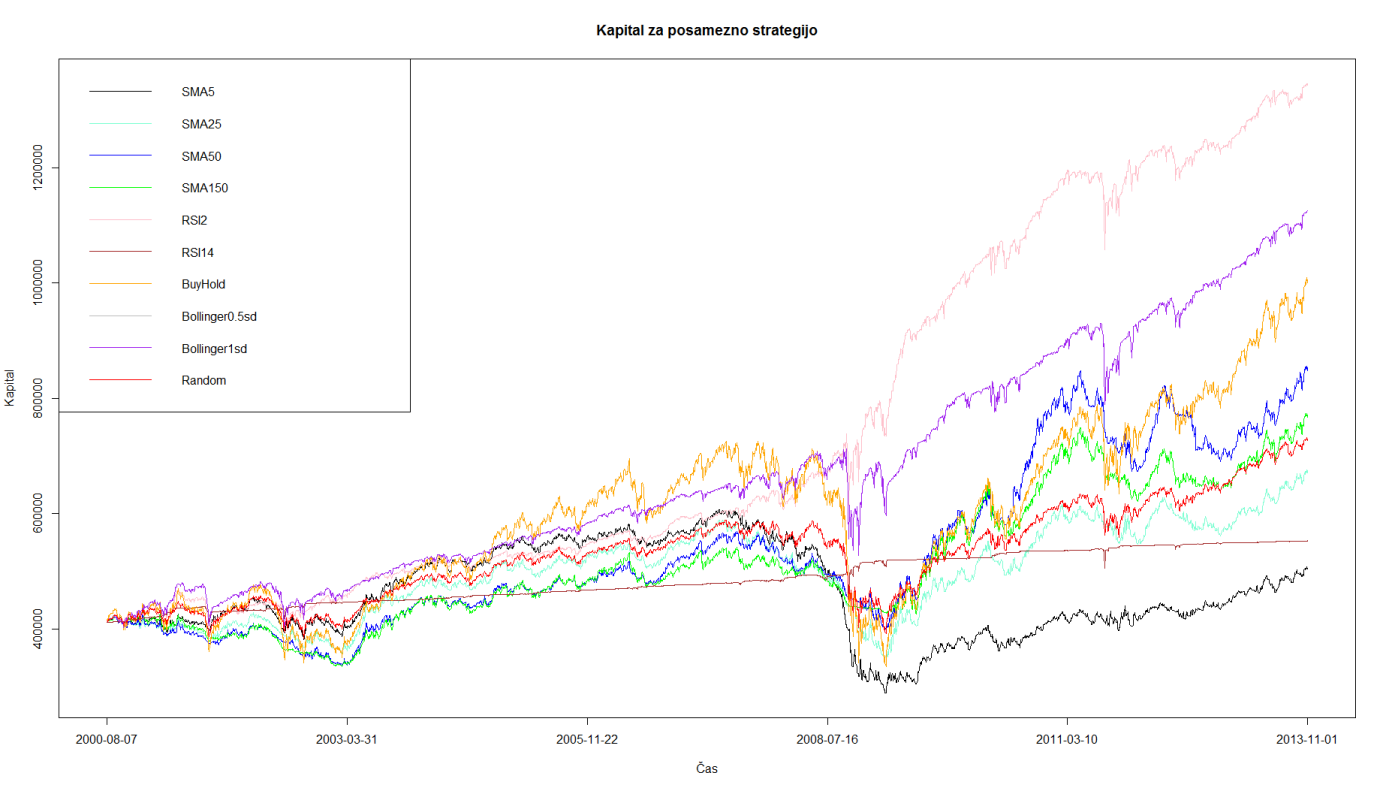
Random: naključno sem izbrala, ali vstopim v trgovalno pozicijo ali ne

Buy&Hold: delnico na začetku kupim in jo imam do konca

**3. uporaba strategij:**

Na vseh delnicah sem uporabila predstavljene strategije. Vsaki delnici v portfelju sem na začetku namenila 1000 eurov kapitala. Ta kapital sem potem uporabila za trgovanje s to delnico. Kapital celotnega portfelja sem izračunala kot vsoto kapitalov za posamezno delnico v portfelju.

Slika prikazuje kapital portfelja za posamezno strategijo tekom 13. let:



Kot je razvidno iz grafa je najboljša strategija za dani portfelj in za dano obdobje strategija RSI14. Sledi ji Boillinger, ki temelji na SMA20 in SD 1 oz. 0.5. na tretjem mestu je pasivna strategija Buy&Hold, kjer na začetku vložimo 1000 eurov v vsako delnico v portfelju in nato ne počnemo nič.

Kot najslabša se je izkazala strategija SMA5. Najmanj volatilna je strategija RSI14, vendar ne ravno donosna.

Opazimo očiten padec vseh strategij v koncu leta 2008, zaradi začetka finančne krize