

Εθνικό Μετσόβιο Πολυτεχνείο

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

**Πρόγραμμα Μεταπτυχιακών Σπουδών Επιστήμης Δεδομένων
και Μηχανικής Μάθησης**



Επίλυση προβλήματος χρηματιστηρίου μέσω Python

Εργασία μαθήματος:

«Προγραμματιστικά Εργαλεία και Τεχνολογίες για την Επιστήμη Δεδομένων»

Βόγκας Ιωάννης, Α.Μ.: 03400206

e-mail: ioannisvogkas@mail.ntua.gr

Ιανουάριος 2024

Εισαγωγή

Σκοπός της παρούσας αναφοράς είναι η κατασκευή ενός αλγορίθμου και η αξιοποίησή του σε μία υποθετική κατάσταση που θα ήταν εφικτή η αγοραπωλησία μετοχών στο χρηματιστήριο σε οποιαδήποτε ημερομηνία στο παρελθόν. Χρησιμοποιώντας ένα σύνολο δεδομένων με τιμές μετοχών από το 1960 μέχρι και το 2017, δημιουργήθηκαν δύο ακολουθίες κινήσεων, διαφορετικού μεγέθους, με σκοπό να αποφέρουν ένα μεγάλο κέρδος.

Προεπεξεργασία Δεδομένων

Το πρώτο στάδιο της διαδικασίας αποτέλεσε η έρευνα των μετοχών για τον εντοπισμό εκείνων που χρησιμοποιήθηκαν στην επίλυση του προβλήματος. Περισσότερες από 7.000 μετοχές ήταν διαθέσιμες σε ένα χρονικό ορίζοντα περίπου 60 ετών, από το 1960 μέχρι το 2017. Η συμμετοχή όλων των μετοχών στην επίλυση του προβλήματος κρίθηκε ανόφελι, αφού πολλές από αυτές αδυνατούσαν να αυξήσουν το κέρδος και καθιστούσαν την διαχείριση τους σημαντικά δυσκολότερη.

Κριτήρια για την επιλογή μιας μετοχής αποτέλεσαν το ποσοστό αύξησης της τιμής της, οι διαθέσιμες ποσότητες για αγορά και πώληση και η ύπαρξη διακυμάνσεων κατά την διάρκεια των ετών. Μετοχές για τις οποίες οι τιμές τους παρουσίαζαν σταθερή ή πτωτική πορεία, αλλά και όσες δεν είχαν μεγάλες διαθέσιμες ποσότητες για αγορά απορρίφθηκαν από την αρχή. Κατόπιν δοκιμών διαφορετικών αλγορίθμων για την έρευνα των επιθυμητών μετοχών διαπιστώθηκε ότι ο συνδιασμός τους με περαιτέρω προσωπική έρευνα για τις πιο κερδοφόρες μετοχές στην πάροδο των χρόνων είχε σημαντικά καλύτερα αποτελέσματα. Ακόμα, η κατασκευή διαγραμμάτων για κάθε μετοχή προσέφερε χρήσιμες πληροφορίες και οδήγησε στην τελική επιλογή των μετοχών.

Καταλήγοντας, χρησιμοποιήθηκαν συνολικά 10 μετοχές στην περίπτωση small, ενώ προστέθηκαν 17 επιπλέον στην περίπτωση large.¹

Κατασκευή αλγορίθμου

Μετά την επιλογή των μετοχών, η επίλυση του προβλήματος συνεχίστηκε με την υλοποίηση του αλγορίθμου για την διαχείριση των δεδομένων. Αρχικά, συγχωνεύτηκαν όλα τα δεδομένα των μετοχών σε μία κοινή δομή δεδομένων και ταξινομήθηκαν σε αυξούσα χρονολογική σειρά.

Στο πρώτο βήμα του αλγορίθμου, ξεκινώντας από την παλαιότερη διαθέσιμη ημερομηνία, γίνεται αναζήτηση μέχρι να βρεθεί η πρώτη ημερομηνία στην οποία είναι εφικτό να γίνει αγορά κάποιας μετοχής. Χωρίς ακόμα να πραγματοποιείται η αγορά αυτή, γίνεται αναζήτηση για την καλύτερη στιγμή που μπορεί να πουληθεί η συγκεκριμένη μετοχή εντός

¹Μετοχές: aapl, nvda, amzn, tsla, googl, goog, mmm, rok, ba, fb (small), wmt, jnj, hd, ma, mo, pg, xom, jpm, pep, cvx, ge, nflx, mcd, msft, ibm, rcl, amd (large).

ένος ορισμένου χρονικού διαστήματος, που ορίζεται διαφορετικά για τις δύο περιπτώσεις και θα αναλυθεί στην συνέχεια.

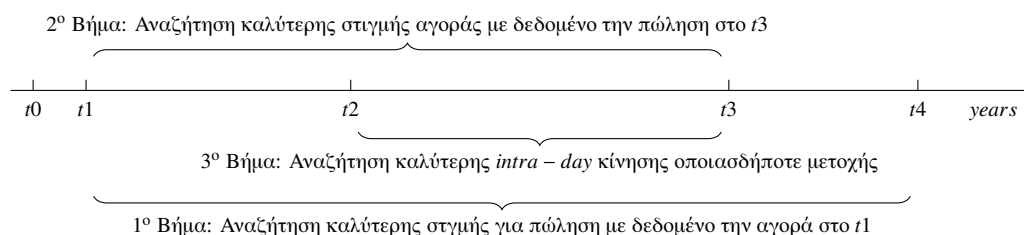
Στο επόμενο βήμα, εξετάζεται αν η πώληση θα είχε αποφέρει μεγαλύτερο κέρδος στην περίπτωση που η αγορά γινόταν σε κάποια μεταγενέστερη ημερομηνία από αυτήν που είχε βρεθεί στην αρχή. Έτσι, έχει αποθηκευτεί η αγορά στο τοπικό ελάχιστο και η πώληση στο τοπικό μέγιστο του διαστήματος αυτού, χωρίς ακόμα να σημαίνει ότι θα πραγματοποιηθεί.

Για την αναζήτηση μίας ακόμα καλύτερης κίνησης στο διάστημα αυτό, εξετάζονται όλες οι πιθανές *intra-day* κινήσεις για όλες τις μετοχές μεταξύ του τοπικού ελαχίστου και του τοπικού μεγίστου που έχουν βρεθεί. Με τον τρόπο αυτό, επιτυγχάνεται η εύρεση μίας αγοραπωλησίας με αρκετά καλό κέρδος στο εξεταζόμενο διάστημα. Οι κινήσεις *intra-day*, εκτός από το γεγονός ότι συνεισφέρουν στην εύρεση μίας καλύτερης επιλογής, προστατεύουν την λειτουργία του αλγορίθμου στην περίπτωση που η πορεία της τιμής της μετοχής που επιλέχθηκε στο συγκεκριμένο διάστημα είναι πτωτική. Συγκρίνοντας το κέρδος της καλύτερης *intra-day* κίνησης με αυτό των τοπικών μεγίστων και ελαχίστων, αποφασίζεται ποιο ζεύγος κινήσεων θα γίνει τελικά. Έτσι, ολοκληρώνεται μία επανάληψη του αλγορίθμου.

Στην συνέχεια, ξεκινώντας από το σημείο που έγινε η τελευταία αγορά (ή η αγοραπωλησία στην *intra-day* περίπτωση) ο αλγόριθμος ψάχνει την κοντινότερη ημερομηνία στην οποία μπορεί να γίνει οποιαδήποτε αγορά και επαναλαμβάνει την διαδικασία. Όταν συναντάει ημερομηνία που έχει ορίσει πώληση, αυξάνει κατάλληλα το διαθέσιμο υπόλοιπο και ξεκινάει εκ νέου την αναζήτηση από την επόμενη μέρα.

Διευκρινίζεται ότι κάθε φορά γίνεται αγορά της μέγιστης δυνατής ποσότητας μίας μετοχής, βάσει των περιορισμών (10% και σειρά κινήσεων) που υπάρχουν και του διαθέσιμου υπολοίπου. Στην περίπτωση *buy-low - sell-high*, μέγιστη ποσότητα επιτρεπόμενης αγοράς ορίζεται το 10% της μικρότερης τιμής όγκου μεταξύ των δύο ημερών.

Τέλος, με την αυθαίρετη επιλογή της αρχικής μετοχής για την οποία γίνεται η αναζήτηση κάθε φορά, ελοχεύει ο κίνδυνος να μην εξεταστεί κάποια άλλη με καλύτερο κέρδος σε αυτό το διάστημα. Όμως, το γεγονός ότι όλες οι μετοχές έχουν αξιολογηθεί ως κερδοφόρες περιορίζει τον κίνδυνο αυτόν και καθιστά τον αλγόριθμο αποδοτικό. Παρακάτω απεικονίζεται η λειτουργία του αλγορίθμου.

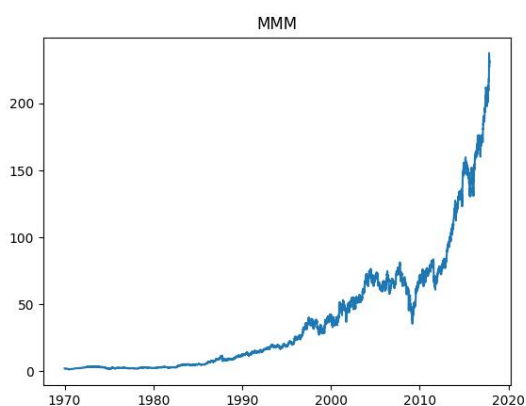


Βήματα:

- Ξεκινώντας από t_0 , αναζητείται η κοντινότερη ημερομηνία για εφικτή αγορά (t_1).

- Με την υπόθεση ότι θα γίνει αγορά στο t_1 , αναζητείται η καλύτερη στιγμή για πώληση στο διάστημα $[t_1, t_4]$.
- Έχοντας βρεί το καλύτερο σημείο για πώληση το t_3 , αναζητείται η χαμηλότερη τιμή αγοράς στο διάστημα $[t_1, t_3]$.
- Έχοντας βρεί τα καλύτερα σημεία αγοράς και πώλησης (t_2, t_3), αναζητείται οποιαδήποτε intra-day κίνηση με μεγαλύτερο κέρδος στο διάστημα $[t_2, t_3]$. Επιλέγεται το ζευγάρι κινήσεων με το μεγαλύτερο κέρδος, μεταξύ των καλύτερων intra-day και buy-low - sell-high κινήσεων.
- Επαναλαμβάνεται η διαδικασία με νέο σημείο αφετηρίας t_0 την νέα ημερομηνία αγοράς.
- Τερματίζεται όταν οι κινήσεις φτάσουν την τιμή μέγιστων κινήσεων N ή όταν εξαντληθούν οι ημερομηνίες.

Εξετάζοντας τα διαγράμματα των μετοχών που επιλέχθηκαν, στην πλειοψηφία τους, παρατηρείται ένα παρόμοιο μοτίβο. Οι τιμές τους παραμένουν σταθερές και σε χαμηλά επίπεδα τα παλαιότερα χρόνια, ενώ παρουσιάζουν αυξομειώσεις τα επόμενα χρόνια. Έτσι, χρειάστηκε να τροποποιηθούν τα διαστήματα, που αναφέρθηκαν παραπάνω, ανάλογα με την χρονολογία στην οποία γίνεται η αναζήτηση. Για τις αναζητήσεις μετά το 1995 ορίζονται μικρότερα διαστήματα, σε σύγκριση με αυτές που γίνονται πριν το 1995, γιατί εκεί λόγω των πολλών διακυμάνσεων είναι πιο εύκολο να βρεθεί ένα ζευγάρι κινήσεων με καλο κέρδος. Ενδεικτικά, αυτό φαίνεται στα Σχήματα 1 και 2.



Σχήμα 1: Τιμή μετοχής MMM

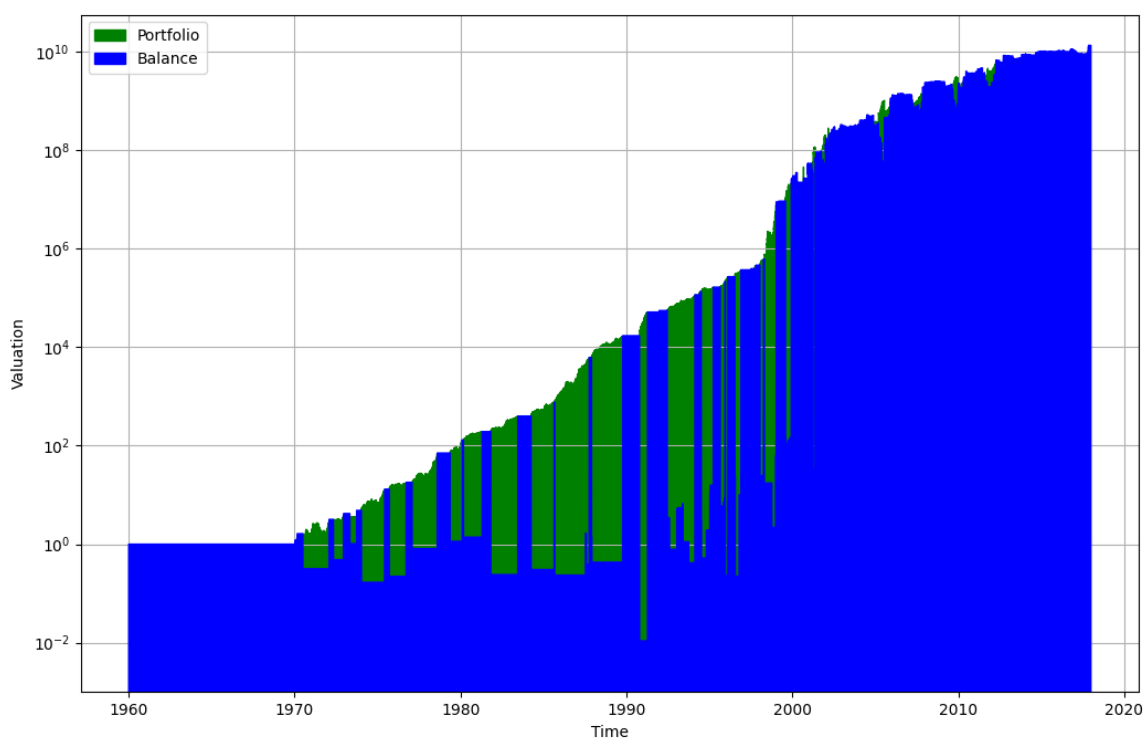


Σχήμα 2: Τιμή μετοχής BA

Περίπτωση small

Για την περίπτωση small, όσο η κάθε πιθανή ημερομηνία αγοράς βρίσκεται πριν από το 1995, το διάστημα στο οποίο αναζητείται το τοπικό μέγιστο περιορίζεται στις 800 ημέρες. Δηλαδή, το διάστημα $[t_1, t_4]$ περιέχει 800 ημέρες, όταν το t_1 βρίσκεται πριν το

1995. Αντίθετα, όταν η πιθανή ημερομηνία αγοράς βρίσκεται μετά το 1995, το διάστημα αναζήτησης γίνεται 300 ημέρες. Λόγω των μεγαλύτερων διακυμάνσεων των τιμών, μετά το 1995, είναι πιο εύκολο να βρεθεί ένα καλό σημείο πώλησης σε σύντομο χρονικό διάστημα και να ξανά γίνει αγορά της ίδιας μετοχής σε μία πιο εύνοική τιμή. Έτσι, ο αλγόριθμος είναι πιο πιθανό να εκμεταλευτεί αυξομειώσεις των τιμών σε σύντομα χρονικά διαστήματα. Επίσης, για να περιοριστούν οι κινήσεις, και να απλουστευθεί ο αλγόριθμος, θεωρήθηκε ότι κάθε μέρα μπορεί να γίνει το πολύ μία αγορά ή μία πώληση ή μία intra-day κίνηση. Η εκτέλεση του αλγορίθμου, σε αυτή την περίπτωση, αποφέρει **13.238.781.835,441563\$**, με **946 κινήσεις** να έχουν πραγματοποιηθεί. Στο Σχήμα 3 φαίνεται η αποτίμηση των μετοχών στην πάροδο των ετών.



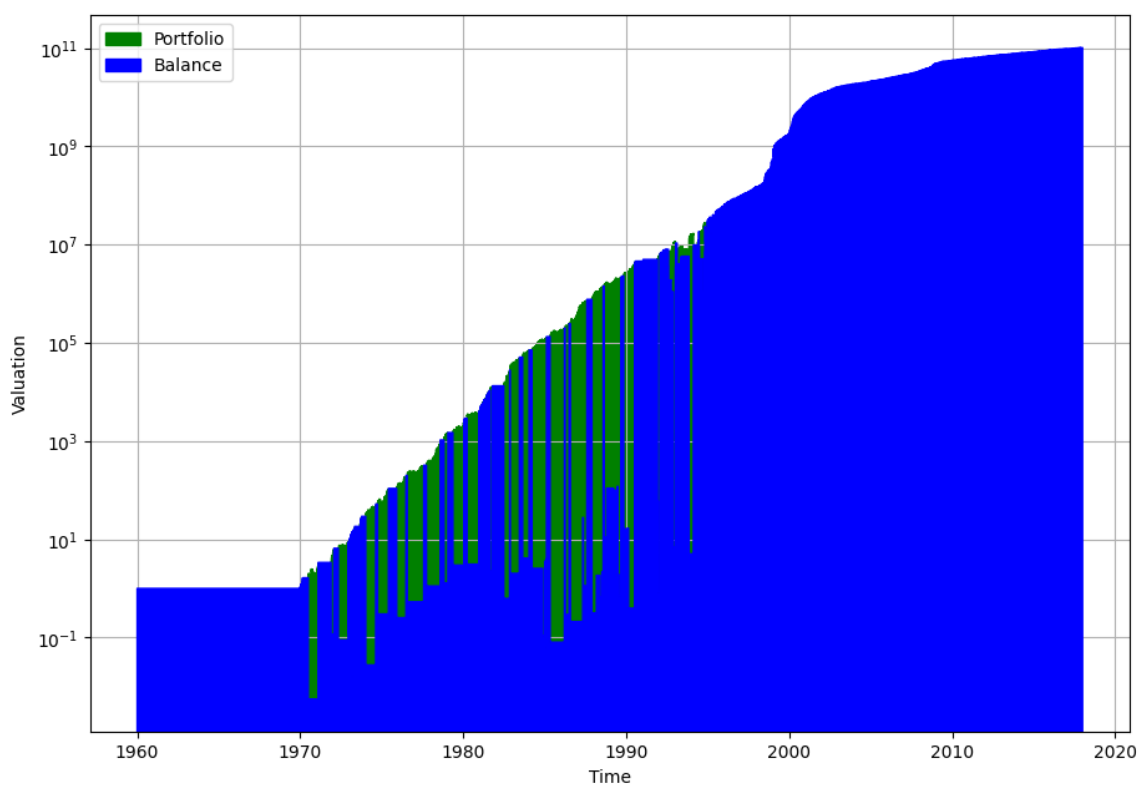
Σχήμα 3: Διάγραμμα αποτίμησης περίπτωσης small

Περίπτωση large

Ο αλγόριθμος φαίνεται να αποδίδει αρκετά καλά για μικρό αριθμό κινήσεων, όμως η ταχύτητα του δεν τον καθιστά κατάλληλο για μεγαλύτερο αριθμό κινήσεων. Έτσι, για την περίπτωση large, δημιουργήθηκε μία αρχική λύση με το συγκεκριμένο αλγόριθμο και στην συνέχεια εμπλουτίστηκε αξιοποιώντας την δυνατότητα των intra-day κινήσεων. Αρχικά, έγινε άρση του περιορισμού σχετικά με τον αριθμό των κινήσεων που μπορούν να γίνουν κάθε μέρα, όπως αυτός αναφέρθηκε στην περίπτωση small. Τα διαστήματα στα οποία

γίνονται αναζητήσεις περιορίστηκαν ως εξής: όσο γίνεται αγορά πριν το 1995, το διάστημα περιέχει 400 ημέρες, ενώ μετά το 1995 περιορίζεται στις 20 μέρες. Σκοπός, όπως και πριν, είναι να βρίσκει ο αλγόριθμος τοπικά μέγιστα και ελάχιστα για να εκμεταλλεύεται αυξομειώσεις των τιμών. Οι διακυμάνσεις των τιμών μετά το 1995 επιτρέπει στον αλγόριθμο να βρίσκει ένα καλό κέρδος σε μικρότερα χρονικά διαστήματα. Όπως αναφέρθηκε και στην ενότητα της προεπεξεργασίας, στην περίπτωση large αξιοποιήθηκαν 27 μετοχές.

Αρχικά, με την χρήση του αλγορίθμου επιτυγχάνεται κέρδος 16.792.480.170,057394\$ με λιγότερες από 10 χιλιάδες κινήσεις. Στο επόμενο βήμα εμπλουτίστηκε η λύση με τις καλύτερες δυνατές, βάσει περιορισμών, intra-day κινήσεις. Δεν εξετάστηκαν κινήσεις πριν το 2000, καθώς θα ήταν δύσκολο να προστεθούν αγορές λόγω χαμηλού υπολοίπου σε πολλά σημεία. Μετά το 2000, ο αλγόριθμος έχει εξασφαλίσει ένα υψηλό υπόλοιπο που καθιστά εφικτό να γίνονται αγορές οποιαδήποτε χρονική στιγμή. Η εκτέλεση του αλγορίθμου, σε αυτή την περίπτωση, αποφέρει **102.120.143.651,0825\$**, με **170.856 κινήσεις** να έχουν πραγματοποιηθεί. Στο Σχήμα 4 φαίνεται η αποτίμηση των μετοχών στην πάροδο των ετών.



Σχήμα 4: Διάγραμμα αποτίμησης περίπτωσης large