



ΙΩΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΠΜΣ “ΨΗΦΙΑΚΕΣ ΕΦΡΜΟΓΕΣ ΚΑΙ ΚΑΙΝΟΤΟΜΙΕΣ”

Ψηφιακές Εφαρμογές για την Αγροτική Παραγωγή και το Περιβάλλον (Digital Applications in Agriculture and Environment)

Σύγκριση Υπολογιστικών Μεθόδων Python–Excel για την Ανάλυση Δεδομένων Θερμοκρασίας
και Υγρασίας

Αναστασία Ρίζου (inf.dig2410)

Inf.dig2410@ionio.gr

Κέρκυρα, Νοέμβριος 2025

| | |
|-------------------|---|
| Εισαγωγή..... | 2 |
| Μεθοδολογία | 2 |
| Ευρήματα | 3 |
| Συμπεράσματα..... | 4 |

Εισαγωγή

Το σύνολο των τεσσάρων ασκήσεων αφορά την ανάπτυξη και αξιολόγηση ενός συστήματος παρακολούθησης περιβαλλοντικών δεδομένων με αισθητήρες θερμοκρασίας και υγρασίας, με στόχο την αξιόπιστη ανάλυση των μετρήσεων και την υποστήριξη αποφάσεων για φαινόμενα ξηρασίας. Στην πρώτη άσκηση δίνεται έμφαση στην επεξεργασία και βαθμονόμηση των

μετρήσεων, στη δεύτερη στην περιγραφική στατιστική ανάλυση, στην τρίτη στη μοντελοποίηση της εξέλιξης ενός δείκτη με λογιστικό μοντέλο, και στην τέταρτη στη διαμόρφωση και αξιολόγηση ενός κανόνα απόφασης (DSS) σε επίπεδο κατηγοριοποίησης επεισοδίων.

Μεθοδολογία

Στην **Άσκηση 1** χρησιμοποιήθηκε χρονοσειρά μετρήσεων θερμοκρασίας και υγρασίας για τον υπολογισμό βασικών περιγραφικών στατιστικών πριν και μετά τον καθαρισμό και τη βαθμονόμηση των δεδομένων. Η θερμοκρασία βαθμονομήθηκε με έναν γραμμικό μετασχηματισμό της μορφής : $\text{temp_cal} = 1.01 * \text{temperature} - 0.2$, όπου ο συντελεστής κλίμακας (gain) 1,01 αυξάνει τις μετρήσεις κατά 1% και ο όρος μετατόπισης (offset) -0.2°C διορθώνει σταθερή απόκλιση του αισθητήρα. Η υγρασία διατηρήθηκε αρχικά χωρίς μετασχηματισμό. Παράλληλα υπολογίσθηκαν μέσοι όροι θερμοκρασίας και υγρασίας στο Excel ώστε να συγκριθούν οι «raw» με τις βαθμονομημένες τιμές.

Στην **Άσκηση 2** εξετάστηκε δείγμα 500 τιμών θερμοκρασίας και υπολογίσθηκαν στο Excel όλα τα βασικά μέτρα κεντρικής τάσης και διασποράς: μέσος όρος, διάμεσος, ελάχιστη και μέγιστη τιμή, διακύμανση και τυπική απόκλιση. Οι τύποι του Excel αξιοποιήθηκαν τόσο σε «raw» μορφή (π.χ. AVERAGE, MEDIAN, VAR.S, STDEV.S) όσο και σε μορφοποιημένο πίνακα αναφοράς, ώστε να αποτυπωθεί με συστηματικό τρόπο η στατιστική εικόνα των θερμοκρασιών.

Στην **Άσκηση 3** χρησιμοποιήθηκε ένα λογιστικό μοντέλο της μορφής:

$$P_{(t)} = \frac{K}{1+ae^{-rt}} + \varepsilon_t ,$$

με αρχική τιμή $P_0 = 15$, ρυθμό αύξησης $r = 0,28$, φερούσα ικανότητα $K = 900$ και προσθήκη θορύβου με τυπική απόκλιση 2,5. Οι τιμές του μοντέλου καταγράφηκαν στη στήλη $P_{(t)}$ για διαδοχικά χρονικά βήματα t , επιτρέποντας τη μελέτη της σύγκλισης του δείκτη προς τη μέγιστη τιμή του και της επίδρασης του θορύβου στο σήμα.

Στην **Άσκηση 4** υλοποιήθηκε σε Excel ο κανόνας απόφασης του συστήματος υποστήριξης αποφάσεων, όπως είχε αρχικά διατυπωθεί σε Python. Στο φύλλο **Sheet1** υπολογίσθηκαν οι δείκτες score_T και score_H από τις βαθμονομημένες θερμοκρασίες και υγρασίες (temperature_cal, humidity_cal), και ο συνολικός πολυκριτηριακός δείκτης score ως συνδυασμός αυτών. Στη συνέχεια ορίστηκε ένας «αυστηρός» κανόνας αναφοράς (y_true), βασισμένος σε θερμοκρασίες $>31^{\circ}\text{C}$, υγρασίες $<38\%$ και χαμηλές τιμές score, καθώς και η πρόβλεψη του συστήματος (y_pred) μέσω της μεταβλητής DSS_alert. Στο φύλλο **Metrics** υπολογίσθηκαν, με χρήση συναρτήσεων SUMPRODUCT και IFERROR, η μήτρα σύγχυσης (TP, FP, FN, TN) και οι μετρικές απόδοσης Precision, Recall, Specificity, Accuracy και F1-score.

Ευρήματα

Στην πρώτη άσκηση, η ανάλυση έδειξε ότι οι μετρήσεις θερμοκρασίας κυμαίνονται ενδεικτικά μεταξύ περίπου 15,6 °C και 29 °C, με μέσο όρο περί τους 22 °C, ενώ η υγρασία κινείται σε μέσο επίπεδο περίπου 55 %. Η εφαρμογή της βαθμονόμησης οδηγεί σε μικρή αλλά συστηματική μετατόπιση του μέσου όρου της θερμοκρασίας, αποτυπώνοντας τη διόρθωση της μεροληψίας του αισθητήρα και βελτιώνοντας τη συγκρισιμότητα των δεδομένων.

Στη δεύτερη άσκηση, το δείγμα των 500 θερμοκρασιών παρουσιάζει μέσο όρο 24,96 °C, διάμεσο 25,01 °C, ελάχιστη τιμή 17,30 °C και μέγιστη 33,74 °C, με τυπική απόκλιση περίπου 2,88 °C. Τα αποτελέσματα αυτά δείχνουν σχετικά συμπαγή κατανομή γύρω από τις τιμές 24–26 °C, με κάποιες υψηλότερες τιμές που αντιστοιχούν σε θερμότερα επεισόδια. Η σύγκριση μέσου και διαμέσου υποδηλώνει σχεδόν συμμετρική κατανομή, χωρίς έντονη ασυμμετρία..

Στην τρίτη άσκηση, το λογιστικό μοντέλο με τις παραμέτρους $P_0 = 15$, $r = 0,28$ και $K = 900$ παράγει μια εξέλιξη του δείκτη $P_{(t)}$ που συγκλίνει προοδευτικά προς τη φέρουσα ικανότητα, με μέση τιμή περί τα 741 και τιμές του άνω τετάρτου τεταρτημορίου πολύ κοντά στο θεωρητικό όριο των 900. Η παρουσία θορύβου προκαλεί παροδικές αποκλίσεις και μηδενισμούς σε ορισμένα χρονικά σημεία, γεγονός που αναδεικνύει τη ρεαλιστική μεταβλητότητα ενός περιβαλλοντικού δείκτη υπό στοχικές διακυμάνσεις.

Στην τέταρτη άσκηση, τα περιγραφικά στατιστικά του φύλλου Metrics δείχνουν μέση βαθμονομημένη θερμοκρασία περίπου 26,23 °C και μέση υγρασία 46,24 %. Η μήτρα σύγχυσης καταγράφει $TP = 67$, $FP = 1$, $FN = 0$ και $TN = 931$, με αντίστοιχες μετρικές $Precision \approx 0,985$, $Recall = 1,0$, $Specificity \approx 0,999$, $Accuracy \approx 0,999$ και $F1 \approx 0,993$. Τα αποτελέσματα αυτά τεκμηριώνουν ότι ο κανόνας DSS εντοπίζει όλα τα πραγματικά επεισόδια ενδιαφέροντος (καμία χαμένη θετική περίπτωση) και παράγει ελάχιστα ψευδώς θετικά, επιτυγχάνοντας εξαιρετικά υψηλή συνολική απόδοση.

Συμπεράσματα

Οι τέσσερις ασκήσεις συγκροτούν μια ολοκληρωμένη αλυσίδα επεξεργασίας και ανάλυσης περιβαλλοντικών δεδομένων αισθητήρων. Στην αρχή, η προσεκτική βαθμονόμηση και καθαρισμός των μετρήσεων διασφαλίζουν ότι τα δεδομένα είναι αξιόπιστα και κατάλληλα για

στατιστική επεξεργασία. Στη συνέχεια, η περιγραφική στατιστική παρέχει σαφή εικόνα της κατανομής και της μεταβλητότητας των θερμοκρασιών, ενώ η λογιστική μοντελοποίηση επιτρέπει τη μελέτη της δυναμικής ενός δείκτη που σχετίζεται με φαινόμενα όπως η ξηρασία. Τέλος, η υλοποίηση και αξιολόγηση ενός κανόνα απόφασης (DSS) στο Excel, σε συμφωνία με την αρχική υλοποίηση σε Python, αποδεικνύει ότι το σύστημα μπορεί να λειτουργήσει ως αξιόπιστο εργαλείο υποστήριξης αποφάσεων, με πολύ υψηλές τιμές ευαισθησίας και ακρίβειας.

Η συνολική εικόνα αναδεικνύει τη σημασία του συνδυασμού προσεκτικής προεπεξεργασίας, στατιστικής ανάλυσης, μοντελοποίησης και τυπικής αξιολόγησης απόδοσης, τόσο για την κατανόηση της συμπεριφοράς των αισθητήρων όσο και για την ασφαλή αξιοποίηση των δεδομένων σε εφαρμογές περιβαλλοντικής παρακολούθησης.