

WATER LEAK AND  
OVERUSE DETECTION  
USING MACHINE  
LEARNING



## Η ΣΗΜΑΣΙΑ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

έλλειψη γλυκού νερού κατά 40% έως το 2030 λόγω της αυξημένης αστικοποίησης, της κλιματικής αλλαγής

ο σύγχρονος τρόπος ζωής είναι επιρρεπής στην υπερκατανάλωση και τη σπατάλη πόρων τόσο περιορισμένων όσο το νερό.

οι διαρροές νερού στα σπίτια μπορεί να περνούν απαρατήρητες για μεγάλα χρονικά διαστήματα πριν ανακαλυφθούν και επισκευαστούν από τους ιδιοκτήτες.

## ΑΙΤΙΑ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Ηλικία του  
σωλήνα

Πίεση νερού

Αλλαγές  
θερμοκρασίας

Κινήσεις  
εδάφους

Σημείωση:

Οι περισσότερες μελέτες επικεντρώνονται σε δίκτυα μεγάλης κλίμακας

# ΠΡΟΣΕΓΓΙΣΕΙΣ

1


Συστήματα ανίχνευσης  
διαρροών νερού με χρήση  
δικτύων αισθητήρων χωρίς  
μηχανική μάθηση

2

Συστήματα ανίχνευσης  
διαρροών νερού με χρήση  
μηχανικής μάθησης.

# ΑΝΙΧΝΕΥΣΗ ΔΙΑΡΡΟΩΝ ΧΩΡΙΣ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (THRESHOLD BASED)

Χρήση Sensors που λειτουργούν με προκαθορισμένα  
όρια (Thresholds)

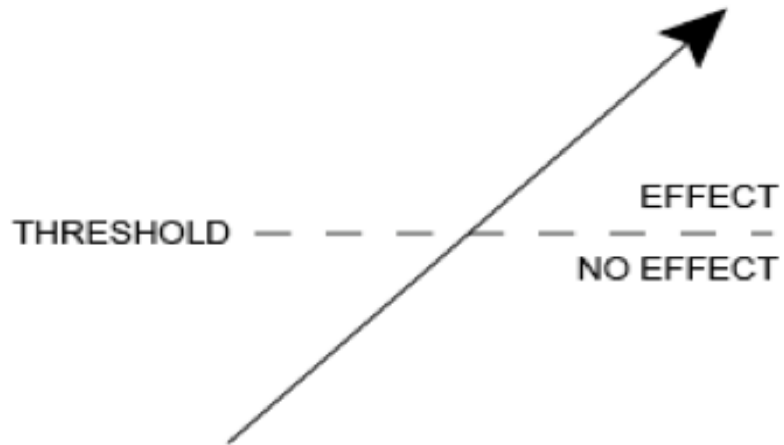


Αν η τιμή ενός Sensor υπερβεί ένα όριο,  
ενεργοποιείται συναγερμός για πιθανή διαρροή

Ειδοποίηση μέσω SMS / Εφαρμογής / LCD οθόνη

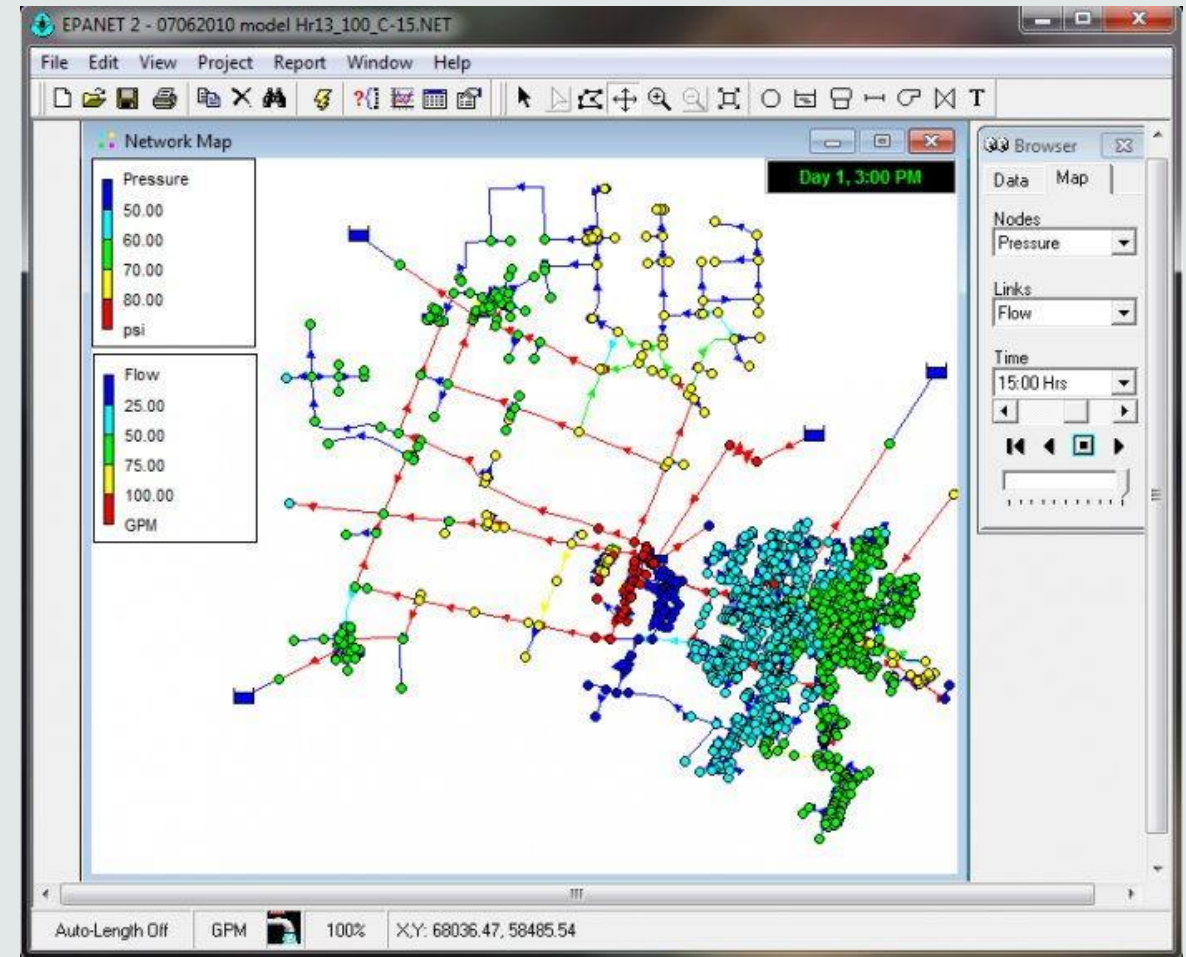
# ΠΕΡΙΟΡΙΣΜΟΙ

- Στατικά Κατώφλια => Μειωμένη ακρίβεια
- Αδυναμία να "μάθουν" από νέα δεδομένα ή να ανιχνεύσουν πιο σύνθετες ανωμαλίες
- Καταλληλότερα για μεγάλα δίκτυα παρά για οικιακή χρήση



# ΑΝΙΧΝΕΥΣΗ ΔΙΑΡΡΟΩΝ ΜΕ ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ (ML)

- Βελτίωση της ακρίβειας και της ευαισθησίας σε σχέση με τις μεθόδους που βασίζονται μόνο σε κατώφλια
- επιβλεπόμενη και μη επιβλεπόμενη μάθηση (Labels)
- Το σύνολο δεδομένων ελήφθη με τη χρήση του λογισμικού **EPANET**



# ΠΡΟΤΕΙΝΟΜΕΝΟ ΣΥΣΤΗΜΑ

Έγκαιρη ανίχνευση διαρροών νερού και υπερβολικής χρήσης νερού σε κατοικίες

Απλό στην εγκατάσταση & Φτηνός εξοπλισμός

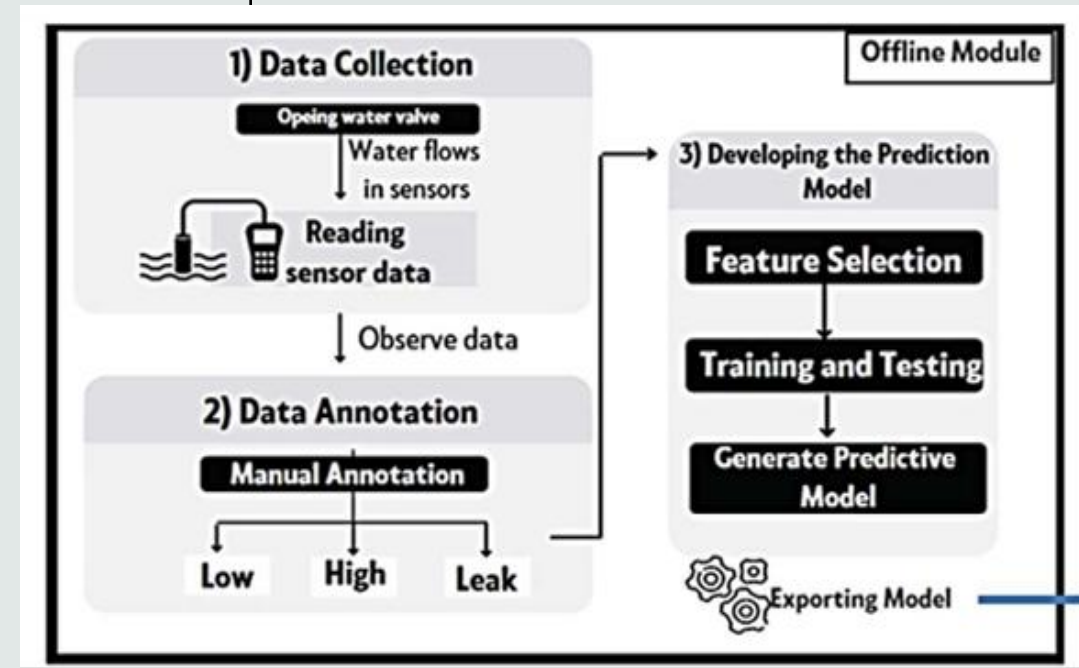
Αξιοποίηση Μηχανικής Μάθησης

Χρήση Random Forest (High Accuracy)



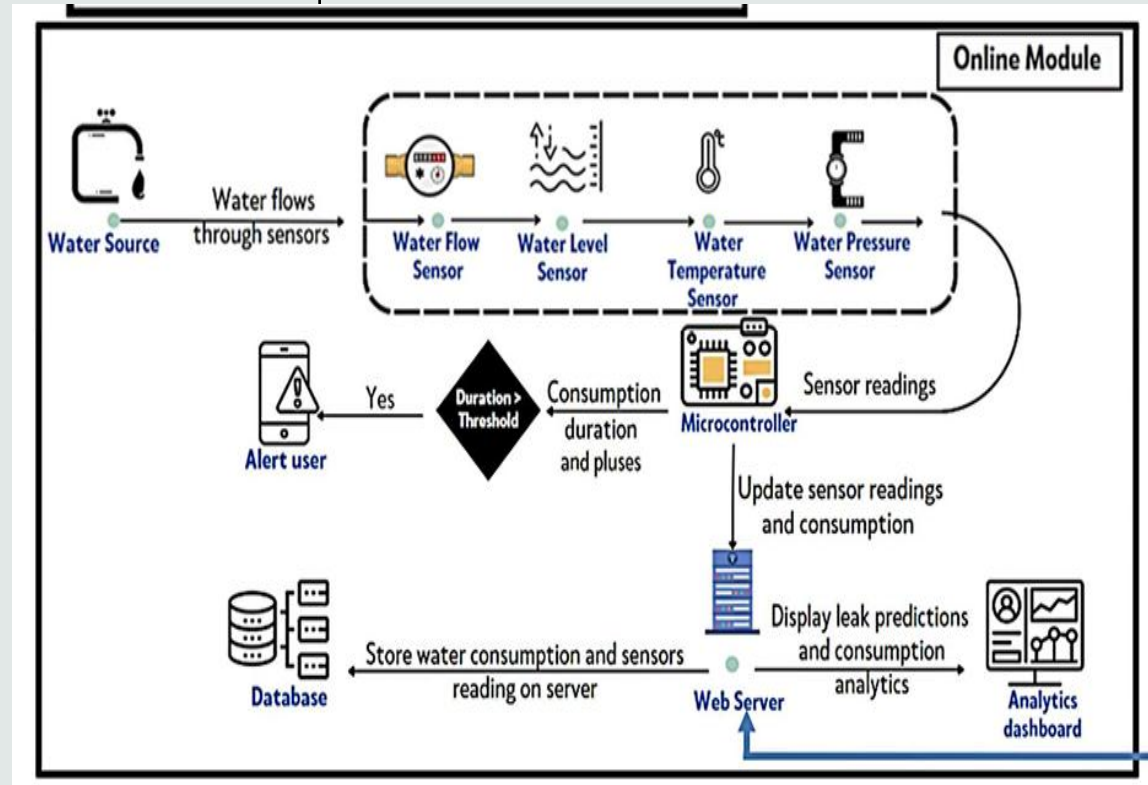
# OFFLINE MODULE

- Έμμεση σύνδεση με το πραγματικό σύστημα ύδρευσης
- Αισθητήρες ροής, πίεσης, στάθμης, θερμοκρασίας (reading data)
- Προσομοίωση σεναρίων: διαρροή, υπερκατανάλωση, κανονική χρήση
- Δημιουργία μοντέλου
- Εξαγωγή του μοντέλου (export model)



# ONLINE MODULE

- Συλλογή δεδομένων από την πηγή
- Σύνδεση με Microcontroller (αποθήκευση / υπολογισμός)
- Σύνδεση με Web Server (& prediction model)
- Πρόβλεψη διαρροής (leak detection)
- Αποστολή ειδοποίησης



# ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

- Καθαρισμός δεδομένων
- Εξάλειψη θορύβου
- Βελτίωση της ποιότητας του σήματος
- Z-score normalization (ίδιο εύρος τιμών)

Table 4 Sensor Readings before Normalization			Table 5 Sensor Readings after Normalization		
PSI	WF	Temp	PSI	WF	Temp
12.072	14072	21.312	-1.516	-0.337	0.553
12.072	17940	21.432	-1.516	0.374	0.874
12.072	16255	21.181	-1.516	0.064	0.203
12.072	18287	21.139	-1.516	0.438	0.09
12.072	14794	21.122	-1.516	-0.205	0.044
12.072	18081	21.113	-1.516	0.4	0.022
12.072	18474	20.944	-1.516	0.472	-0.432
12.072	14834	20.775	-1.516	-0.197	-0.885
12.072	18814	20.741	-1.516	0.535	-0.976
12.072	14905	21.908	-1.516	-0.184	2.149
12.072	18495	21.312	-1.516	0.476	0.553
12.072	18607	21.147	-1.516	0.497	0.112
12.072	18625	20.893	-1.516	0.5	-0.568
12.072	16344	21.187	-1.516	0.08	0.218

# ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ

- Decision Tree
- K-NN
- Naïve Bayse
- Random Forest

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

# ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΑΞΙΝΟΜΗΣΗΣ

- Διαρροή: 4% από το σύνολο των δεδομένων εκπαίδευσης

Μοντέλα	Σύγκριση κατά τάξη			
		ΔΙΑΡΡΟΗ	ΧΑΜΗΛΗ	HIGH
Δέντρο αποφάσε ων	PR	66%	84%	84%
	Rec	66%	90%	74%
	Acc	83%		
K-NN	PR	75%	87%	84%
	Rec	65%	90%	80%
	Acc	86%		
Naïve Bayes	PR	72%	91%	55%
	Rec	46%	80%	88%
	Acc	79%		
Τυχαίο δάσος	PR	<u>75%</u>	<u>89%</u>	<u>83%</u>
	Rec	<u>83%</u>	<u>90%</u>	<u>79%</u>
	Acc	<u>87%</u>		

\*Ακρίβεια = PR, Ανάκληση = Rec, Ακρίβεια = Acc

# ΣΥΜΠΕΡΑΣΜΑ

- Οι περισσότερες μελέτες εστιάζουν σε δίκτυα μεγάλης κλίμακας (όχι οικιακά)
- Η οικιακή διαρροή νερού μπορεί να περάσει απαρατήρητη (=> μεγάλες απώλειες)

## i) Ακρίβεια κατά κατηγορία (75%)

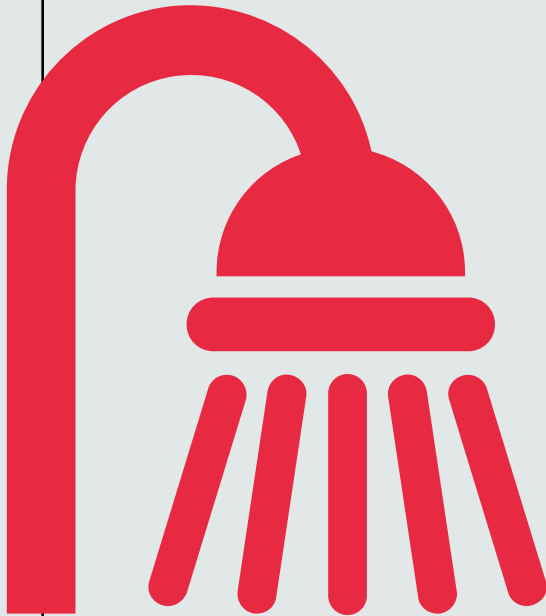
Το ποσοστό των σωστών προβλέψεων του μοντέλου για κάθε ξεχωριστή κατηγορία (π.χ., διαρροή, υπερβολική χρήση).

## ii) Συνολική ακρίβεια (87%)

Το ποσοστό όλων των σωστών προβλέψεων (ανεξαρτήτως κατηγορίας) σε σχέση με το σύνολο των δοκιμών.

## iii) Ανάκληση (83%)

Η **ανάκληση** (recall) δείχνει πόσο καλά το μοντέλο εντοπίζει όλες τις περιπτώσεις διαρροής (ακόμα και τις πιο δύσκολες).



## ΕΥΧΑΡΙΣΤΩ ΓΙΑ ΤΗΝ ΠΡΟΣΟΧΗ ΣΑΣ

- Άρθρο: Smart Residential Water Leak and Overuse Detection System Using Machine Learning
- Heba Ismail\*, Rawan Elabyad\*, Arwa Dyab