**A Machine learning approach for forest fires.**

**Περίληψη** (Ελληνικά)

Ο σημαίνων ρόλος της *Τεχνητής* *Νοημοσύνης*, πάνω στην πρόβλεψη των φυσικών καταστροφών, έχει δημιουργήσει αισιόδοξες προοπτικές, για ένα νέο πιο ασφαλές μέλλον. Συγκεκριμένα, μέσω της *Μηχανικής* *Μάθησης* – Υπολογιστικής Μάθησης, η οποία, είναι μέρος της *Τεχνητής Νοημοσύνης*∙ δύναται, να αξιοποιηθούν στατιστικά μοντέλα, ούτως ώστε να δημιουργηθούν:

* Μοντέλα πρόβλεψης δασικών πυρκαγιών,
* Μοντέλα πρόβλεψης διάρκειας δασικών πυρκαγιών.

Ειδικότερα, έχουν προταθεί διάφοροι αλγόριθμοι, Μηχανικής Μάθησης, για αυτό το σκοπό, όπως: Continuous Convolutional Neural Network (CCNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), Convolutional Neural Network (CNN), K – nearest neighbor (KNN), XGBoost και Encoder – Based CNN. .

Εν’ συνέχεια, αυτής της δημοσίευσης, θα εξεταστούν οι προαναφερόμενες μεθόδους.

**Εισαγωγή**

[***Parallel SVM model for forest fire prediction***](https://www.sciencedirect.com/science/article/pii/S2666222121000046)***.***

Εξ’ ανατολής, και δη στην Ινδία, το 2021, διεξήχθη, η εξής έρευνα: *Parallel SVM model for forest fire prediction*[[1].](#_Kajol_R_Singh,) Η συγκεκριμένη έρευνα αναδεικνύει, τις προκλήσεις που αντιμετωπίζουν, με συγκεκριμένους αλγόριθμους, αλλά και την τεχνική επίλυσης αυτών, των εμποδίων. Ο αλγόριθμος *Convolutional Neural Network* (CNN) επιδεικνύει πρόβλημα υπερβολικής προσαρμογής στην εκπαίδευση. Οι μέθοδοι των *Polynomial Neural Network* (PNN) και *SVM*, έχουν προτιμήσεις προς κλάσεις με μεγάλο αριθμό κλάσεων, και είναι λιγότερο αποτελεσματικές σε μεγάλα σύνολα δεδομένων [[1].](#_Kajol_R_Singh,) Εν’ προκειμένου να προσπεραστούν οι δυσκολίες, χρησιμοποιήθηκε, ένα παράλληλο *SVM* μοντέλο, το οποίο, διασπάστηκε σε πολλά υπό – σύνολα, π.χ *libSVM*, *light SVM* κ.α., ούτως ώστε να εκτελέσει ταυτόχρονα πολλαπλές επεξεργασίες και σε μετεωρολογικά δεδομένα.

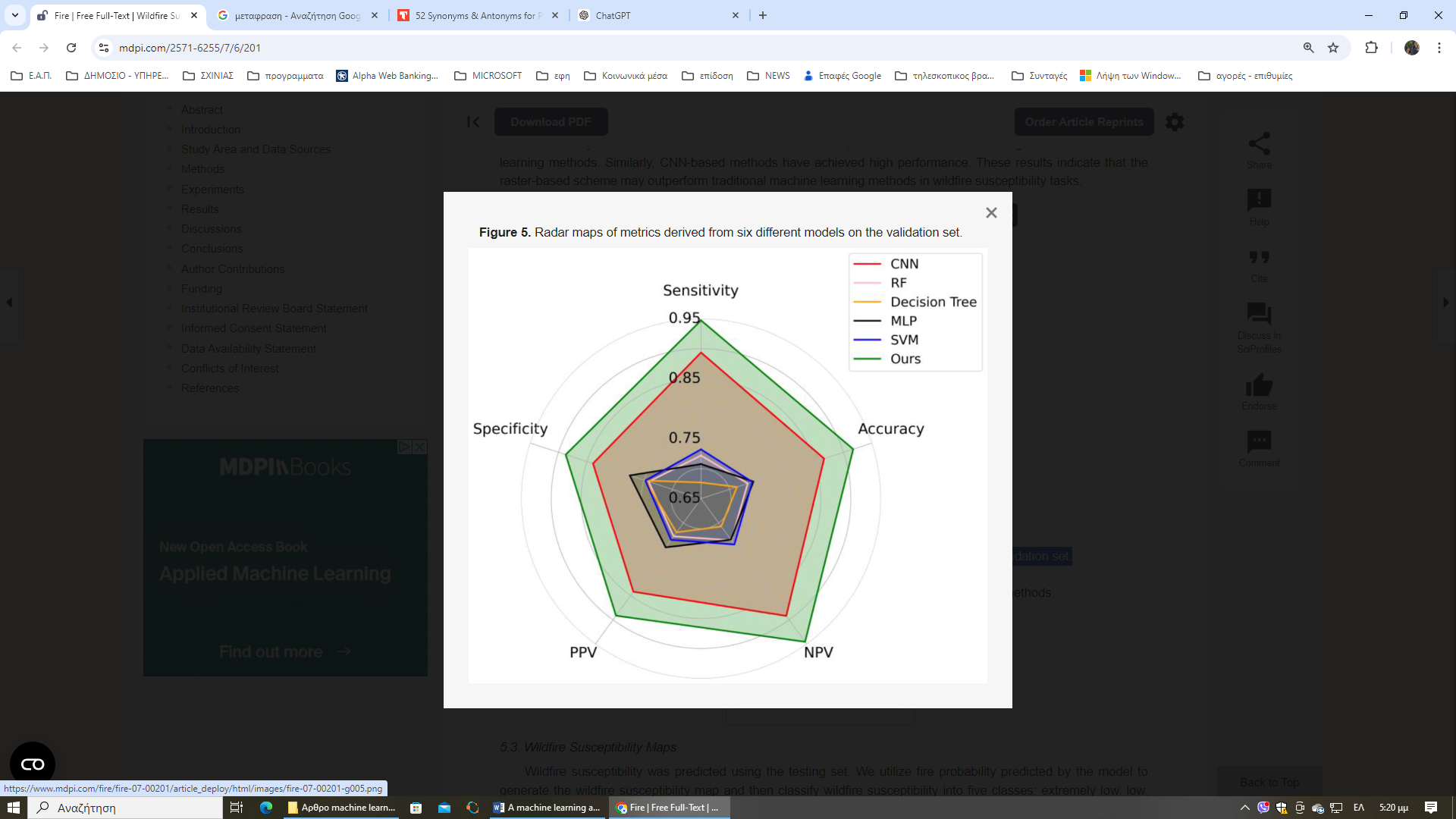
Ο δείκτης Root Mean Square Error (RMSE) έδειξε ότι ο αλγόριθμος *Parallel SVM* είχε τα λιγότερα σφάλματα, σε σχέση με άλλους αλγόριθμους [[1].](#_Kajol_R_Singh,)

Σύγκριση του προτεινόμενου *Parallel SVM* με υπάρχοντα μοντέλα πρόβλεψης δασικών πυρκαγιών.

| **Methods** | **RMSE** | **MSE** | **MAE** |
| --- | --- | --- | --- |
| CCN | 68.3 | 4665.5 | 673.9 |
| MPNN | 63.8 | 4076.4 | 652.2 |
| PNN | 63.7 | 4069.3 | 466 |
| RBF | 68.1 | 4638.2 | 887.9 |
| SVM | 63.5 | 4042 | 502 |
| Parallel SVM | 63.45 | 4026 | 453 |

[***Wildfire Susceptibility Prediction on a CA – Based CCNN with Active Learning Optimization.***](https://www.mdpi.com/2571-6255/7/6/201)

Από την Κίνα, το 2024, δημοσιεύτηκε, το άρθρο, στο MDPI: *Wildfire Susceptibility Prediction on a CA – Based CCNN with Active Learning Optimization* [[2].](#_Qiuping_Yu,_Yaqin" \t "_parent) Η εν’ λόγω, μελέτη έχει ως γνώμονα, ότι οι παραδοσιακοί αλγόριθμοι, πρόβλεψης πυρκαγιών, όπως: *Random* *Forest* (RF), *Logistic* *Regression* (LR), *Support* *Vector* *Machines* (SVM), *Multilayer* *Perceptron* *Neural* *Network* (MPNN), και *Artificial* *Neural* *Network* (ANN), εστιάζουν, σε μία περιοχή, με συγκεκριμένα χαρακτηριστικά. Ως λύση, για να υπερκεραστούν τα εμπόδια, χρησιμοποιήθηκε ο αλγόριθμος *Continuous Convolutional Neural Network (*CCNN) σε συνδυασμό με γεωγραφικές συντεταγμένες – *Coordinate Attention* (CA), και μετεωρολογικά δεδομένα. Η ακρίβεια του συγκεκριμένου μοντέλου, ανήλθε στο 91,7%, ενώ η τιμή Area Under the Curve (AUC) έφθασε στο 0,9487 [[2].](#_Qiuping_Yu,_Yaqin)

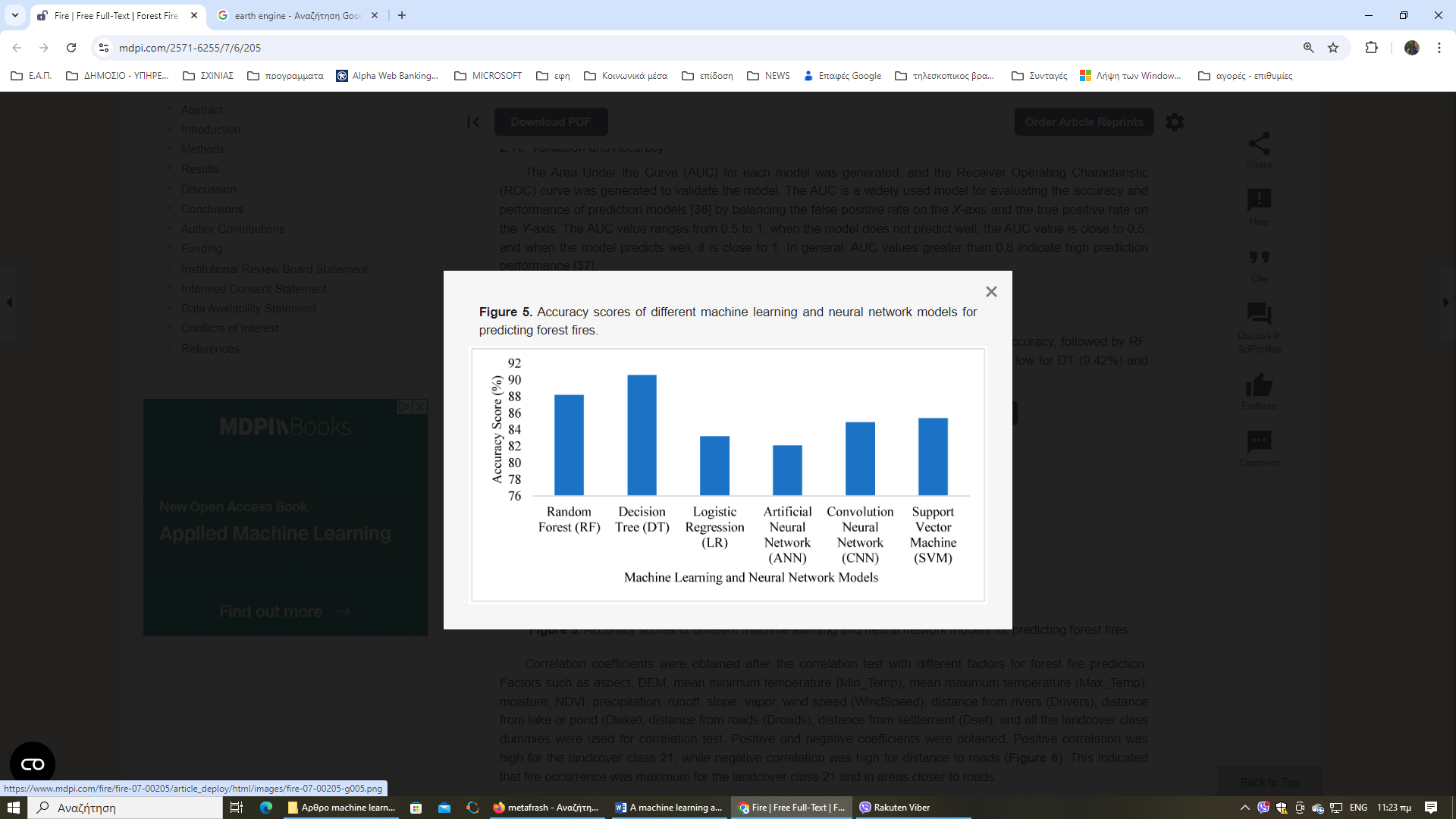
Χάρτης ραντάρ μετρήσεων από έξι διαφορετικά μοντέλα στην επικύρωση.

[***Forest Fire Prediction: A Spatial Machine Learning and Neural Network***.](https://www.mdpi.com/2571-6255/7/6/205)

Από την Νότιο Καρολίνα, της Αμερικής, το 2024, παρουσιάστηκε, στο MDPI το άρθρο, με την έρευνα: *Forest Fire Prediction: A Spatial Machine Learning and Neural Network* [[3].](#_Sanjeev_Sharma,_Puskar) Η συγκεκριμένη έρευνα, κινήθηκε προς τους κλασικούς αλγόριθμους πρόβλεψης, όπως: *Decision Tree* (DT*), Random Forest* (RF), *Logistic Regression* (LR), *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), και *Convolutional* *Neural Network* (CNN). Για τις ανάγκες, της έρευνας, χρησιμοποιήθηκαν Δορυφορικά δεδομένα (Earth Engine, Copernicus, Nasa), στοιχεία πυρκαγιών, και μετεωρολογικά δεδομένα.

Όλα τα προαναφερόμενα μοντέλα, πέτυχαν ακρίβεια πρόβλεψης άνω του 80%. Παρόλα ταύτα, o αλγόριθμος *Decision Tree* (DT), είχε ανακριβή πρόβλεψη, πυρκαγιάς, σε σύγκριση με το correlation test, όταν χρησιμοποιεί τους συντελεστές συσχέτισης. Τα μοντέλα, *Artificial Neural Network* (ANN), και *Convolutional* *Neural Network* (CNN), ναι μεν διαχειρίστηκαν μεγάλο όγκο δεδομένων ωστόσο η ακρίβεια τους, ήταν χαμηλή. Ο αλγόριθμος *Support Vector Machine* (SVM), παρουσίασε υψηλή ικανότητα πρόβλεψης αλλά είχε και ελλείψεις. Αντίστοιχα, ο *Logistic Regression* (LR), παρότι χρησιμοποιείται σε μελέτες με γεωγραφικά δεδομένα GIS, δεν απέδωσε μία ακριβή πρόβλεψη [[3].](#_Sanjeev_Sharma,_Puskar)

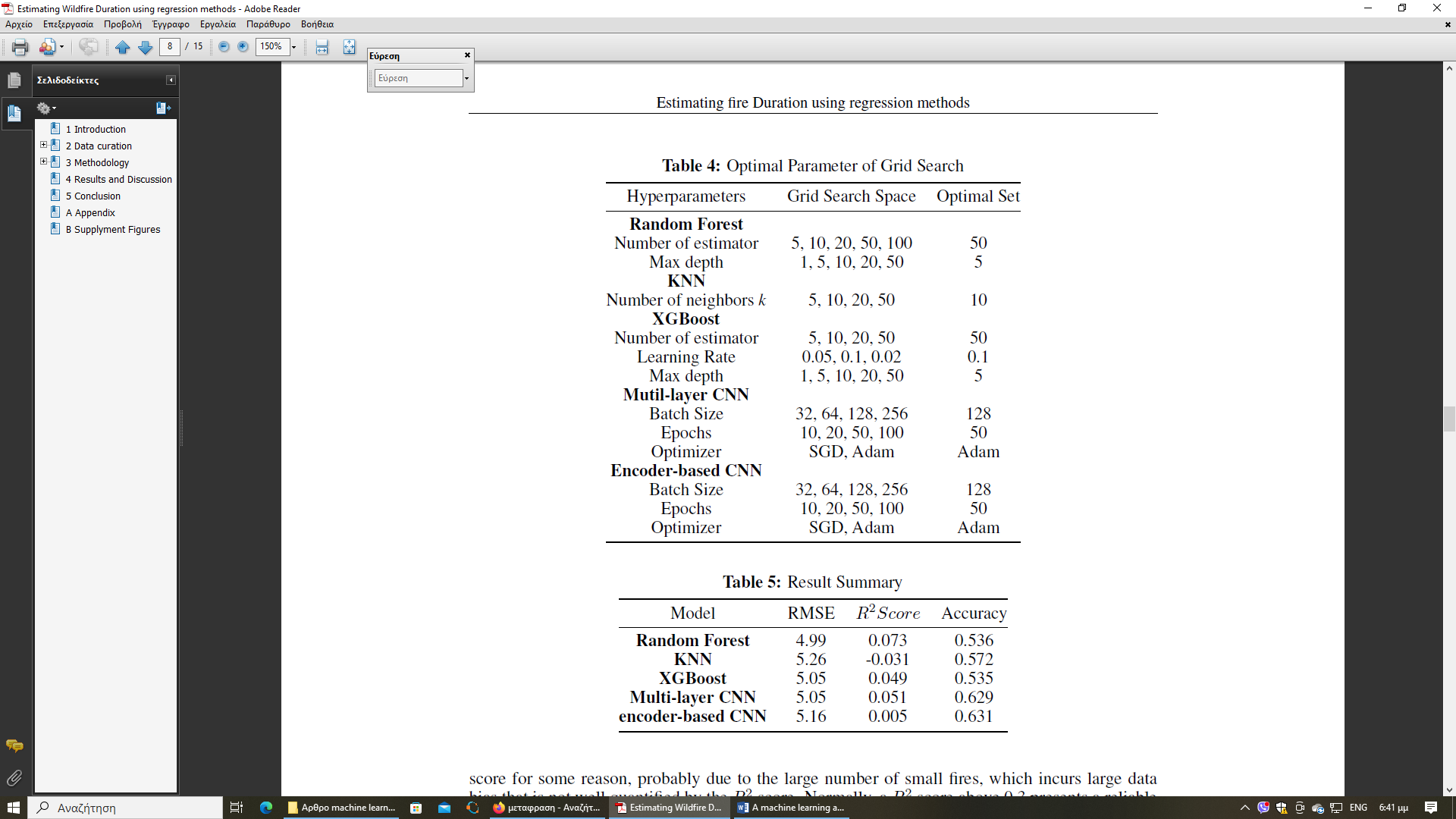
Το συμπέρασμα αυτών, ήταν ότι το correlation test δείχνει την σωστή θεωρητική κατεύθυνση, σε σχέση με άλλες μεθόδους, οι οποίες βασίζονται στη σημασία των χαρακτηριστικών [[3].](#_Sanjeev_Sharma,_Puskar)



[***Estimating Wildfire Duration using regression methods***](https://arxiv.org/pdf/2308.08936v1)

Στο συνέδριο Molecular Machine Learning Conference 2023, παρουσιάστηκε, η έρευνα: *Estimating Wildfire Duration using regression methods* [[4].](#_Hansong_Xiao._Estimating) Αυτή η εργασία, προβάλλει τις δυνατότητες, της πρόβλεψης, της χρονικής διάρκειας μίας πυρκαγιάς. Γι’ αυτό το σκοπό, τα μοντέλα ML, χωρίστηκαν σε δύο κατηγορίες. Αυτά που διαχειρίστηκαν τα δεδομένα: *Random Forest* (RF), *K* – nearest neighbor (KNN), *XGBoost*, και στα μοντέλα που διαχειρίζονται εικόνες, όπως το: *Convolutional* *Neural Network* (CNN), και το *Encoder – Based CNN.* Οι πληροφορίες για τα δεδομένα, για τους αλγόριθμους, προήλθαν από: ιστορικά στοιχεία πυρκαγιών, δορυφορικά και μετεωρολογικά στοιχεία [[4].](#_Hansong_Xiao._Estimating)

Το μοντέλο, εν προκειμένου να έχει ακριβείς προβλέψεις, χρειάζεται μεγάλο όγκο δεδομένων, ούτως ώστε να αποκτήσει λογικό βάρος και μετατόπιση (weight and bias). Η πρόβλεψη ακρίβειας του, κυμαίνεται άνω του 80%, στις πυρκαγιές, οι οποίες, έχουν διάρκεια άνω των 10 ημερών. Ολοκληρώνοντας, να σημανθεί ότι το μοντέλο δεν μπορεί να επιτύχει καλή βαθμολογία στο δείκτη R2, συνεπώς χρειάζεται ένας άλλος στατιστικός συντελεστής, για να αντικαταστήσει το R2 [[4].](#_Hansong_Xiao._Estimating)



**Abstract (Αγγλικά)**

The important role of *Artificial Intelligence* about the prediction on natural disasters, creates optimistic prospects for a new safer future. Specifically, through *Machine Learning*, which is one part of *Artificial Intelligence*, there are several statistical models, that can be used, to create:

* Models for predicting forest - wildfires.
* Models for predicting the duration of forest - wildfires.

In particular, various *Machine learning* algorithms have been proposed in this interpretation, such as: *Continuous Convolutional Neural Network* (CCNN), *Support* *Vector Machine* (SVM), *Decision Tree* (DT), *Random Forest* (RF), *Logistic Regression* (LR), *Artificial Neural Network* (ANN), *Convolutional Neural Network* (CNN), *K – nearest neighbor* (KNN), *XGBoost*.

Consequently, in this publication, the aforementioned methods will be considered.

**Introduction**

[***Parallel SVM model for forest fire prediction.***](https://www.sciencedirect.com/science/article/pii/S2666222121000046)

From the east, specifically in India, in 2021, was conducted the following research: *Parallel SVM model for forest fire prediction* [[1].](#_Kajol_R_Singh,) This specific research highlights the challenges, they face, with specific algorithms, but also refers to the technique of solving these obstacles. The algorithms *Convolutional Neural Network* (CNN) has a problem of over – fitting in training. The *Polynomial Neural Network* (PNN) and *SVM* methods, shows predilection toward classes with a large number of instances and less efficient on large datasets. In order to overcome the difficulties, a parallel *SVM* model was used, which split into several subsets, e.g: *libSVM*, *light SVM*, etc., in sequence to simultaneously multiple processing’s on meteorological data as well.

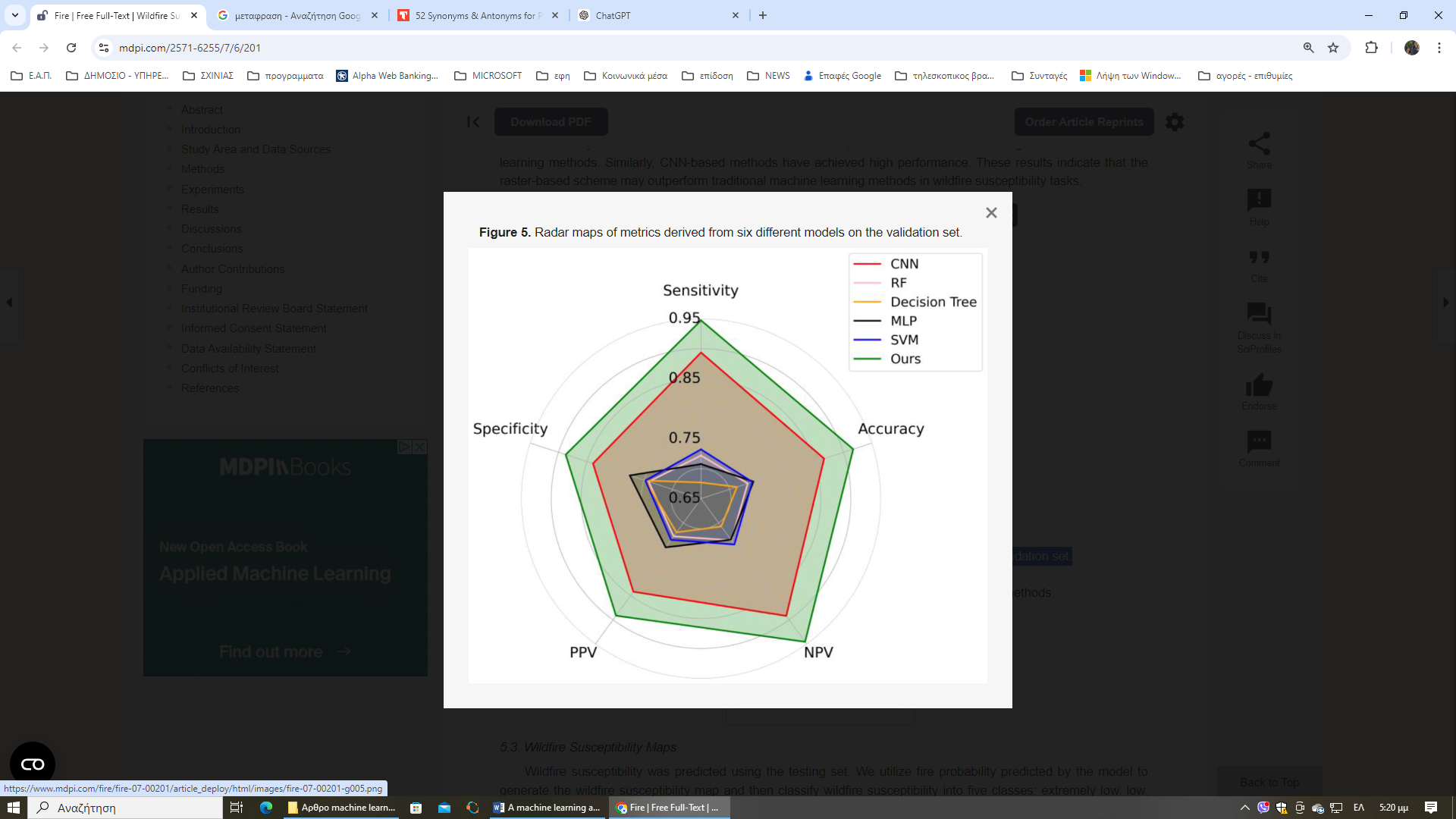
The Root Mean Square Error (RMSE) index showed that the *Parallel SVM* algorithm had the least errors compared to other algorithms [[1].](#_Kajol_R_Singh,)

Comparison of the proposed *parallel SVM* with existing models for Forest Fire prediction.

| **Methods** | **RMSE** | **MSE** | **MAE** |
| --- | --- | --- | --- |
| CCN | 68.3 | 4665.5 | 673.9 |
| MPNN | 63.8 | 4076.4 | 652.2 |
| PNN | 63.7 | 4069.3 | 466 |
| RBF | 68.1 | 4638.2 | 887.9 |
| SVM | 63.5 | 4042 | 502 |
| Parallel SVM | 63.45 | 4026 | 453 |

[***Wildfire Susceptibility Prediction on a CA – Based CCNN with Active Learning Optimization.***](https://www.mdpi.com/2571-6255/7/6/201)

From China, 2024, the article: *Wildfire Susceptibility Prediction on a CA – Based CCNN with Active Learning Optimization****,*** was published on MDPI [[2].](#_Qiuping_Yu,_Yaqin) This scholarship, was based on the fact, that traditional fire prediction algorithms, such as: *Random* *Forest* (RF), *Logistic* *Regression* (LR), *Support* *Vector* *Machines* (SVM), *Multilayer* *Perceptron* *Neural* *Network* (MPNN), and *Artificial* *Neural* *Network* (ANN), focus on one area, with specific characteristics. As a way to overthrow the obstacles, was selected the *Continuous Convolutional Neural Network (*CCNN) combined with geographical *Coordinate Attention* (CA) and meteorological data. The accuracy of the specific model reached 91.7%, while the Area Under the Curve (AUC) value reached 0.9487 [[2].](#_Qiuping_Yu,_Yaqin)

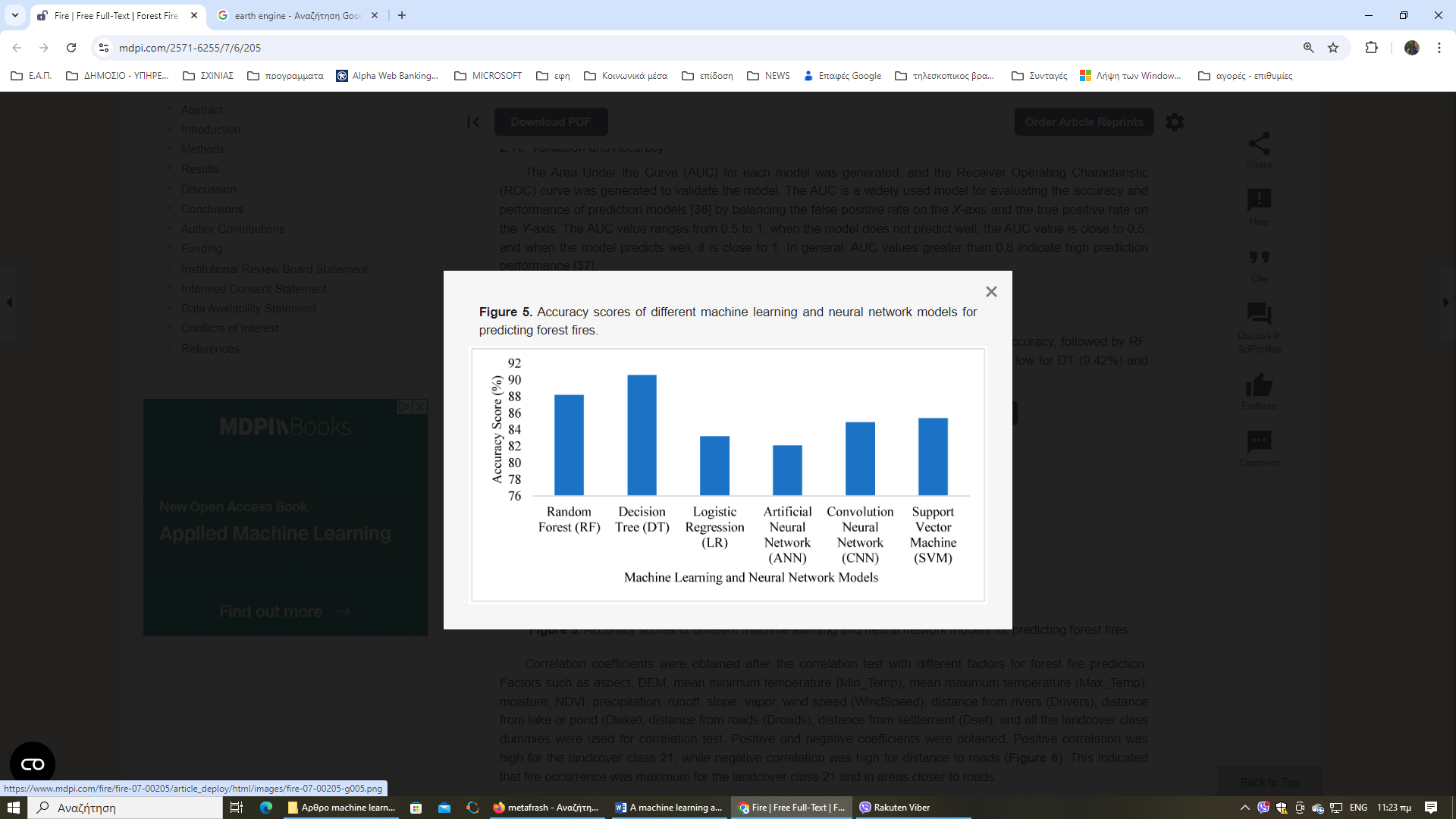
Radar maps of metrics derived from six different models on the validation set.

[***Forest Fire Prediction: A Spatial Machine Learning and Neural Network***.](https://www.mdpi.com/2571-6255/7/6/205)

From South Carolina, America, 2024, the MDPI article was presented, on the research: *Forest Fire Prediction: A Spatial Machine Learning and Neural Network* [3]. This research, moved to the classic prediction algorithms, such as: *Decision Tree* (DT*), Random Forest* (RF), *Logistic Regression* (LR), *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), και *Convolutional* *Neural Network* (CNN). For the research, they used, satellite data (Earth Engine, Copernicus, Nasa), fire and meteorological data [[3].](#_Sanjeev_Sharma,_Puskar)

The aforementioned models achieved prediction accuracies exceeding 80%. Though, the algorithm *Decision Tree* (DT*),* was inaccurate to the hazard, compared to thecorrelation test when utilizes the correlation coefficients*.* The models, *Artificial Neural Network* (ANN), and *Convolutional* *Neural Network* (CNN), could handle large amount of data, but thein accuracy was low. The *Support Vector Machine* (SVM) algorithm, had a high predictive ability, but also had shortcomings. Respectively, the *Logistic Regression* (LR), which is mainly used in studies with geographical data GIS, did not obtain a precise prediction [[3].](#_Sanjeev_Sharma,_Puskar)

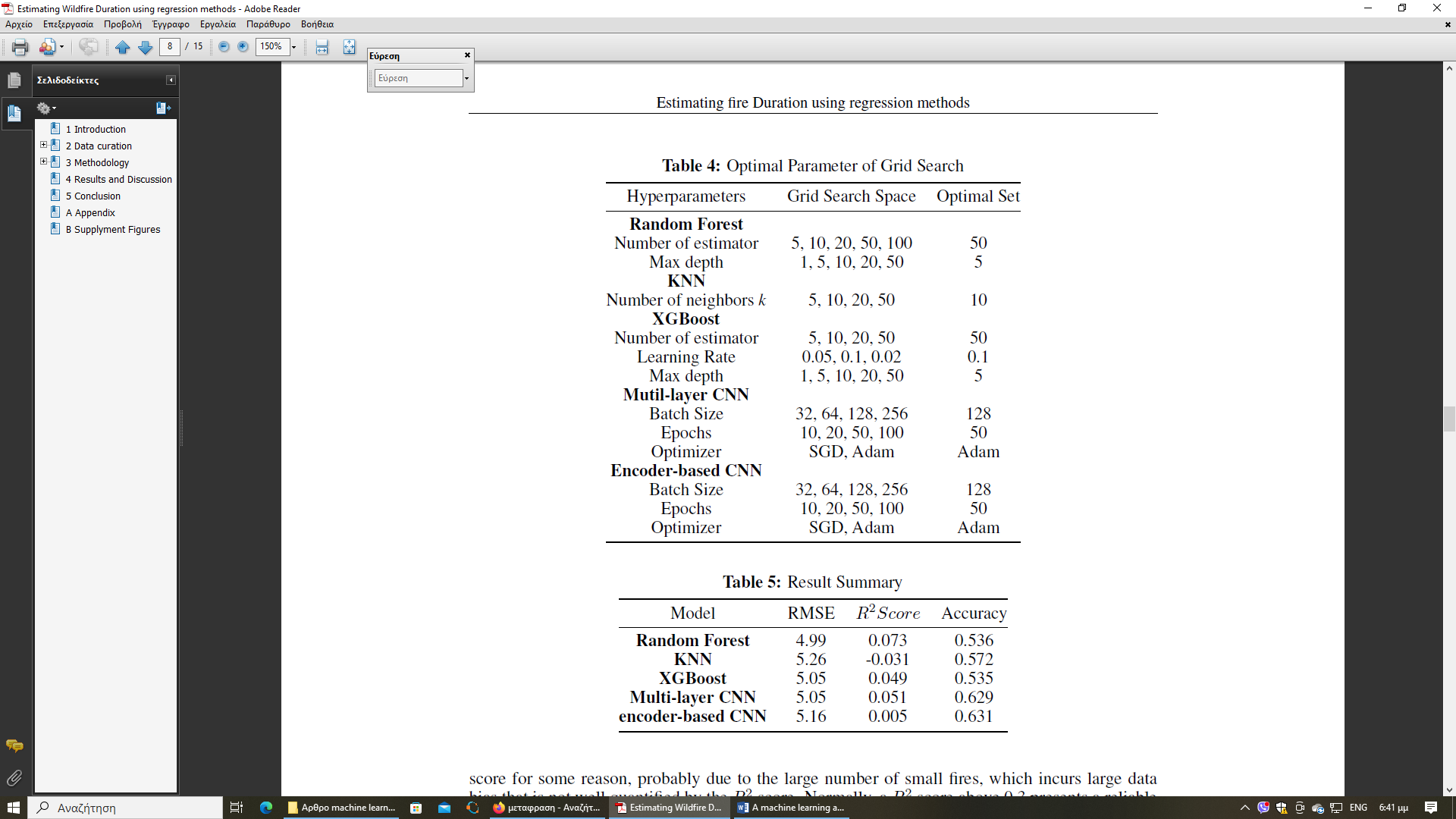
Their outcome, was that the correlation tests shows the right theoretical direction, in relation to other methods, which are based on the importance of the characteristics [[3].](#_Sanjeev_Sharma,_Puskar)



[***Estimating Wildfire Duration using regression methods***](https://arxiv.org/pdf/2308.08936v1)

On the conference, Molecular Machine Learning Conference 2023, it was presented the survey: *Estimating Wildfire Duration using regression methods* [[4].](#_Hansong_Xiao._Estimating) This paper, highlights the possibilities of predicting the fire duration. For this purpose, the ML models were split into two parts. Those that managed the data: *Random Forest* (RF), *K* – nearest neighbor (KNN), *XGBoost*, and the image - based model, such as: *Convolutional* *Neural Network* (CNN), και το *Encoder – Based CNN.* The information, for the algorithms data, came from: historical fire data, satellite and meteorological data [[4].](#_Hansong_Xiao._Estimating)

In order to make accurate predictions, the model, requires a large amount of input data so that to learn a reasonable weight and bias. It’s accuracy forecast ranges over 80%, with fires, which have a duration more than 10 days. In conclusion, it should be noted that the model cannot achieve a good score on the R2 index, so it is needed another coefficient statistic to replace R2 [[4].](#_Hansong_Xiao._Estimating)



**References**

# [Kajol R Singh, K.P. Neethu, K Madhurekaa, A harita, Pushpa Mahan. Parallel SVM model for forest fire prediction]( Kajol R Singh, K.P. Neethu, K Madhurekaa, A harita, Pushpa Mahan. Parallel SVM model for forest fire predictionSoft Computing Letters, volume 3, December 2021.Retrieved from: https://doi.org/10.1016/j.socl.2021.100014)

[Soft Computing Letters, volume 3, December 2021.]( Kajol R Singh, K.P. Neethu, K Madhurekaa, A harita, Pushpa Mahan. Parallel   SVM model for forest fire predictionSoft Computing Letters, volume 3, December 2021.Retrieved from: https://doi.org/10.1016/j.socl.2021.100014)

[Retrieved from: https://doi.org/10.1016/j.socl.2021.100014]( Kajol R Singh, K.P. Neethu, K Madhurekaa, A harita, Pushpa Mahan. Parallel   SVM model for forest fire predictionSoft Computing Letters, volume 3, December 2021.Retrieved from: https://doi.org/10.1016/j.socl.2021.100014)

* 2. <https://www.sciencedirect.com/science/article/pii/S2666222121000046>

# 

# [Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire](Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization. Fire 2024, Volume 7 (6), pp. 201.Retrieved from: https://doi.org/10.3390/fire7060201https://www.mdpi.com/2571-6255/7/6/201) *[Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization.](Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization. Fire 2024, Volume 7 (6), pp. 201.Retrieved from: https://doi.org/10.3390/fire7060201https://www.mdpi.com/2571-6255/7/6/201)* [Fire 2024, Volume 7 (6), pp. 201.](Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization. Fire 2024, Volume 7 (6), pp. 201.Retrieved from: https://doi.org/10.3390/fire7060201https://www.mdpi.com/2571-6255/7/6/201)

[Retrieved from: https://doi.org/10.3390/fire7060201](Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization. Fire 2024, Volume 7 (6), pp. 201.Retrieved from: https://doi.org/10.3390/fire7060201https://www.mdpi.com/2571-6255/7/6/201)

[https://www.mdpi.com/2571-6255/7/6/201](Qiuping Yu, Yaqin Zhao, Zixuan Yin, Zhihao Xu. Wildfire Susceptibility Prediction Based on a CA-Based CCNN with Active Learning Optimization. Fire 2024, Volume 7 (6), pp. 201.Retrieved from: https://doi.org/10.3390/fire7060201https://www.mdpi.com/2571-6255/7/6/201)

# Sanjeev Sharma, Puskar Khanal. Forest Fire Prediction: *A Spatial Machine Learning and Neural Network Approach.*

Fire 2024, volume 7 (6), pp. 205.

Retrieved from: <https://doi.org/10.3390/fire7060205>

<https://www.mdpi.com/2571-6255/7/6/205>

# [Hansong Xiao. Estimating Wildfire Duration using regression methods. arXiv:2308.08936v1. Molecular Machine Learning Conference 2023. Retrieved from: https://arxiv.org/pdf/2308.08936v1](Hansong Xiao. Estimating Wildfire Duration using regression methods. arXiv:2308.08936v1. Molecular Machine Learning Conference 2023. Retrieved from: https://arxiv.org/pdf/2308.08936v1)