ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΕΧΝΟΛΟΓΙΕΣ ΚΙΝΗΤΟΥ ΥΠΟΛΟΓΙΣΜΟΥ ΜΕ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ
ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2024-2025

2Η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ
«ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΓΙΑ ΑΝΑΠΤΥΞΗ ΣΕ ΣΥΣΚΕΥΕΣ ΜΕ ΠΕΡΙΟΡΙΣΜΕΝΟΥΣ ΠΟΡΟΥΣ»

ΕΠΙΚΟΙΝΩΝΙΑ: mcml_lab2@icbnet.ntua.gr

ΕΙΣΑΓΩΓΗ

Αντικείμενο της εργαστηριακής άσκησης είναι η ενσωμάτωση ενός νευρωνικού δικτύου σε έναν 8-bit μικροελεγκτή (MCU), ο οποίος είναι εξοπλισμένος με προκαθορισμένες προδιαγραφές υλικού, για την επίλυση του προβλήματος της ανίχνευσης δικτυακών εισβολών σε ένα περιβάλλον IoT.

Οι μικροελεγκτές παρουσιάζουν από τη φύση τους εξαιρετικά περιορισμένες υπολογιστικές και αποθηκευτικές δυνατότητες, γεγονός που καθιστά αναγκαία τη βελτιστοποίηση των βαθιών νευρωνικών δικτύων πριν από την ανάπτυξή τους σε αυτούς. Το πρόβλημα της ανίχνευσης εισβολών, από την άλλη, έχει αρκετές ιδιαιτερότητες, με πιο βασική την ανισορροπία των κλάσεων, αφού κάθε δικτυακή επίθεση παράγει διαφορετικό όγκο κίνησης.

Το τελικό μοντέλο που θα παραχθεί θα πρέπει να ικανοποιεί: (α) τους στόχους επίδοσης που ορίζει η εφαρμογή, όπως είναι η ακρίβεια ταξινόμησης, και (β) τους περιορισμούς που επιβάλλει το υλικό του μικροελεγκτή, όπως είναι η ανάγκη για 8-bit κβαντοποίηση και οι περιορισμοί στη μνήμη και στη χωρητικότητα αποθήκευσης.

Λόγω όλων των παραπάνω, η απόδοση ενός μοντέλου αξιολογείται χρησιμοποιώντας μια μετρική που λαμβάνει υπόψιν την ακρίβεια ταξινόμησης (top-1 accuracy), το F1-score και το πλήθος των παραμέτρων του μοντέλου. Όσο υψηλότερη είναι η ακρίβεια και το F1-score και όσο λιγότερες είναι οι παράμετροι, τόσο υψηλότερη θα είναι και η μετρική αξιολόγησης.

Η παρούσα εργαστηριακή άσκηση παρέχει μια ολοκληρωμένη ευκαιρία μάθησης, εξοπλίζοντας τους σπουδαστές με την τεχνογνωσία που απαιτείται για την πλοήγηση στο περίπλοκο πεδίο της προσαρμογής και βελτιστοποίησης βαθιών νευρωνικών δικτύων για την εκτέλεση σε περιβάλλοντα περιορισμένων πόρων. Υπογραμμίζει τη σημασία της επίτευξης μιας λεπτής ισορροπίας μεταξύ της αποτελεσματικότητας και της υπολογιστικής ικανότητας του μοντέλου, προσφέροντας πρακτικές γνώσεις για προβλήματα του πραγματικού κόσμου.

Η εκτέλεση της άσκησης θα πραγματοποιηθεί ανά ομάδες, όπως αυτές έχουν ήδη σχηματιστεί για την πρώτη εργαστηριακή άσκηση.

ΔΕΔΟΜΕΝΑ

Το σύνολο δεδομένων που θα χρησιμοποιηθεί είναι ένα υποσύνολο του CICIoT2023[1], ένα σύνολο δεδομένων που παρουσιάστηκε το 2023 από το Canadian Institute for Cybersecurity (CIC) και έκτοτε χρησιμοποιείται ευρέως για προβλήματα ανίχνευσης δικτυακών εισβολών.

Οι δημιουργοί του dataset έχουν εξάγει 46 χαρακτηριστικά από τις δικτυακές ροές, τα οποία μπορούν να χρησιμοποιηθούν για την εκπαίδευση αλγορίθμων ML/DL. Το αρχικό σύνολο δεδομένων περιέχει πάνω από 45 εκατομμύρια δείγματα σε 34 κλάσεις, με μια κλάση να αντιστοιχεί στην καλοήθη (benign) κίνηση

[1] https://www.unb.ca/cic/datasets/iotdataset-2023.html

και οι υπόλοιπες 33 να αφορούν διάφορες δικτυακές επιθέσεις (attacks). Οι 33 επιθέσεις μπορούν να ομαδοποιηθούν σε 7 κατηγορίες, ώστε τελικά να προκύψει ένα πρόβλημα ταξινόμησης 8 κλάσεων.

Το πλήθος δειγμάτων ανά κλάση τόσο του αρχικού όσο και του υποδειγματοληπτημένου συνόλου φαίνονται στον παρακάτω πίνακα. Η υποδειγματοληψία έγινε με τέτοιον τρόπο, ώστε τα δείγματα κάθε κατηγορίας να μην ξεπερνούν τα 150 χιλιάδες, οδηγώντας τελικά σε περίπου 940 χιλιάδες δείγματα συνολικά, για αυτό και ονομάζουμε το νέο σύνολο δεδομένων CICIoT2023-1M.

ΑΑ	**Κλάση**	**CICIoT2023**	**CICIoT2023-1M**
1	DDoS (Distributed Denial of Service)	33.984.560	150.000
2	DoS (Denial of Service)	8.090.738	150.000
3	Mirai	2.634.124	150.000
4	Benign	1.098.195	150.000
5	Spoofing	486.504	150.000
6	Recon (Reconnaissance)	354.565	150.000
7	Web	24.829	24.829
8	BruteForce	13.064	13.064
-	ΣΥΝΟΛΟ	46.686.579	937.893

Κάθε δείγμα (γραμμή) στο σύνολο δεδομένων αφορά μια ροή (flow). Μια ροή αντιπροσωπεύει μια ακολουθία πακέτων IP που μοιράζονται ένα σύνολο κοινών χαρακτηριστικών, που συνήθως ρέουν μεταξύ της ίδιας πηγής και προορισμού μέσα σε ένα συγκεκριμένο χρονικό πλαίσιο. Τυπικά ορίζεται από έναν συνδυασμό πεδίων στις επικεφαλίδες των πακέτων IP, που συχνά αποκαλούνται "5-πλειάδες" (5-tuples) στα δίκτυα IP:

1. **Διεύθυνση IP πηγής:** Προσδιορίζει τον αποστολέα του πακέτου.
2. **Διεύθυνση IP προορισμού:** Προσδιορίζει τον παραλήπτη του πακέτου.
3. **Θύρα προέλευσης:** Προσδιορίζει τον αριθμό θύρας στο μηχάνημα προέλευσης (για πρωτόκολλα όπως το TCP και το UDP).
4. **Θύρα προορισμού:** Προσδιορίζει τον αριθμό θύρας στο μηχάνημα προορισμού (για πρωτόκολλα όπως το TCP και το UDP).
5. **Πρωτόκολλο:** Υποδεικνύει το πρωτόκολλο του επιπέδου μεταφοράς (π.χ. TCP, UDP, ICMP).

Μια ροή μπορεί να περιλαμβάνει πρόσθετα πεδία αλλά η 5-πλειάδα είναι συνήθως επαρκής για τον μοναδικό προσδιορισμό μιας ροής.

ΜΙΚΡΟΕΛΕΓΚΤΗΣ

Θεωρούμε ότι ο MCU στον οποίο επιθυμούμε να αναπτύξουμε το μοντέλο ανίχνευσης εισβολών είναι ο STM32F091RC[2] από τη σειρά STM32F0 της STMicroelectronics, η οποία βασίζεται στον πυρήνα ARM Cortex-M0. Ο συγκεκριμένος πυρήνας έχει σχεδιαστεί ώστε να είναι ένας χαμηλής κατανάλωσης και οικονομικά αποδοτικός επεξεργαστής, οπότε δεν διαθέτει αποκλειστικό υλικό για υπολογισμούς κινητής υποδιαστολής (floating-point unit, FPU). Επομένως, έχει τη δυνατότητα να εκτελεί μόνο ακέραια κβαντισμένα μοντέλα.

Τα χαρακτηριστικά του MCU που μας ενδιαφέρουν είναι ο αποθηκευτικός του χώρος (flash memory) και η στατική μνήμη τυχαίας προσπέλασης (static random-access memory, SRAM):

- Flash Memory: **128 KB**
- SRAM: **32 KB**

ΜΟΝΤΕΛΟ

Το απλούστερο νευρωνικό δίκτυο που μπορεί να κατασκευαστεί για την επίλυση της συγκεκριμένης διεργασίας είναι ένα Multi-Layer Perceptron (MLP), στο οποίο πρέπει να καθοριστούν κατ' ελάχιστο το

[2] https://www.st.com/en/microcontrollers-microprocessors/stm32f091rc.html

πλήθος των κρυφών επιπέδων, το πλήθος των νευρώνων ανά κρυφό επίπεδο και η συνάρτηση ενεργοποίησης των κρυφών επιπέδων.

Στο εισαγωγικό εργαστήριο κατασκευάστηκε ένα τέτοιο δίκτυο με 20 κρυφά επίπεδα, 64 νευρώνες ανά επίπεδο και συνάρτηση ενεργοποίησης ReLU. Η συγκεκριμένη αρχιτεκτονική περιέχει 82.568 παραμέτρους και μετά από κατάλληλη προεπεξεργασία δεδομένων και εκπαίδευση πετυχαίνει ακρίβεια περίπου 93% και F1-score περίπου 87%.

ΚΡΙΤΗΡΙΑ ΕΠΙΔΟΣΗΣ

Θεωρούμε ότι το αποτύπωμα μνήμης του μοντέλου θα είναι τουλάχιστον όσο είναι το μέγεθός του, καθώς θα πρέπει όλες οι παράμετροι να βρίσκονται στη RAM κατά την εκτέλεση. Για αυτόν τον λόγο, θεωρούμε ότι **το μεγαλύτερο μοντέλο που μπορεί να διαχειριστεί ο μικροελεγκτής δεν θα πρέπει να ξεπερνά τις 10.000 παραμέτρους**, που οδηγούν σε αποτύπωμα μνήμης τουλάχιστον 9.8 KB, δηλαδή περίπου το 1/3 της SRAM.

Επιπλέον, θεωρούμε πως **η ακρίβεια του μοντέλου δεν πρέπει να πέσει κάτω από 90%**.

ΒΗΜΑΤΑ ΕΚΤΕΛΕΣΗΣ ΑΣΚΗΣΗΣ

(Α) Το πρώτο βήμα αφορά την **προεπεξεργασία των δεδομένων**.

Τα απαραίτητα βήματα προεπεξεργασίας που πρέπει να εφαρμοστούν είναι:

- Διαχωρισμός του συνόλου εκπαίδευσης στα υποσύνολα εκπαίδευσης (train) και επικύρωσης (val). Μπορείτε να δοκιμάσετε διάφορα ποσοστά διαχωρισμού.
- Κανονικοποίηση (normalization). Μπορείτε να δοκιμάσετε διάφορες τεχνικές. Αν η μέθοδος κανονικοποίησης που θα εφαρμόσετε απαιτεί προσαρμογή στα δεδομένα, τότε αυτή η προσαρμογή πρέπει να γίνει ***μόνο*** *στο σύνολο εκπαίδευσης (μετά και τον διαχωρισμό του συνόλου επικύρωσης)*. Παράδειγμα αποτελεί η κανονικοποίηση z-score, όπου απαιτείται η εκμάθηση της μέσης τιμής και της τυπικής απόκλισης κάθε χαρακτηριστικού στο σύνολο εκπαίδευσης, προτού μετασχηματίσει τα δεδομένα.

Τα προαιρετικά βήματα προεπεξεργασίας που μπορούν να εφαρμοστούν είναι:

- Επαύξηση του συνόλου εκπαίδευσης για την αντιμετώπιση της ανισορροπίας των κλάσεων.
- Μηχανική χαρακτηριστικών (εξαγωγή/επιλογή).

(Β) Το δεύτερο βήμα αφορά την **κατασκευή** και την **εκπαίδευση του μοντέλου**.

Υπάρχουν οι εξής περιορισμοί:

- Τα επίπεδα του μοντέλου θα πρέπει να επεξεργάζονται την είσοδο σειριακά (sequentially). Επομένως, δεν μπορούν να υπάρχουν παράλληλοι κλάδοι (branches) μέσα στο δίκτυο. Για αυτές τις περιπτώσεις ενδείκνυται το Sequential API του TensorFlow.
- Δεν μπορούν να χρησιμοποιηθούν όλα τα επίπεδα που είναι διαθέσιμα στο TensorFlow λόγω των περιορισμών της κβαντοποίησης. Πρέπει να φροντίσετε ότι θα μπορεί να εφαρμοστεί στο μοντέλο επίγνωση κβαντοποίησης.

Επιπλέον, δεν υπάρχει περιορισμός ως προς τις τιμές των υπερπαραμέτρων εκπαίδευσης. Μπορείτε να δοκιμάσετε διάφορα μεγέθη δέσμης, διάφορους βελτιστοποιητές, διάφορες τιμές για τις υπερπαραμέτρους αυτών (π.χ. ρυθμός μάθησης), κ.α.

(Γ) Το τρίτο βήμα αφορά τη **βελτιστοποίηση του μοντέλου** με την εφαρμογή **επίγνωσης κβαντοποίησης** και την επανεκπαίδευσή του.

Βρείτε και εξηγήστε ποιος είναι ο λόγος ύπαρξης των μη εκπαιδεύσιμων παραμέτρων κατά την εφαρμογή της επίγνωσης κβαντοποίησης.

Αναλύστε το πλήθος και τη λειτουργία τους.

(Δ) Διαγωνισμός Kaggle

Αφού έχετε διεξάγει έναν αριθμό από πειράματα, μπορείτε να υποβάλετε τις προβλέψεις των βέλτιστων μοντέλων σας στο Kaggle ώστε να αξιολογηθούν ως προς τα τρία κριτήρια ενδιαφέροντος: την ακρίβεια ταξινόμησης, το F1-score και το μέγεθος.

Στον παρακάτω σύνδεσμο θα βρείτε τον διαγωνισμό, τα δεδομένα, καθώς και το notebook «MCML 24-25 LAB2 MAIN», στο οποίο θα πρέπει να βασιστείτε για την εκπόνηση της άσκησης:

https://www.kaggle.com/competitions/mcml-24-25-lab-2-intrusion-detection-in-iot

ΠΑΡΑΔΟΤΕΑ

Η παράδοση της άσκησης περιλαμβάνει τα ακόλουθα:

- Το python notebook σε μορφή .ipynb με όνομα «notebook_#*TEAM*.ipynb». Φροντίστε τα κελιά να έχουν εκτελεστεί / να διαθέτουν έξοδο (output) ώστε να φαίνονται οι έξοδοι των συναρτήσεων, η εκπαίδευση, οι καμπύλες εκπαίδευσης, κ.λπ. Το notebook αποτελεί ταυτόχρονα την αναφορά σας. Συμπληρώστε στα markdown κελιά *TODO* τις απαντήσεις σας, είτε αφορούν κώδικα, είτε επεξήγηση.
- Το τελικό μοντέλο στη μορφή .tflite με όνομα «model_int_#*TEAM*.tflite».

#*TEAM* είναι το όνομα της ομάδας σας (π.χ. 01, 13, A, C). Συγκεντρώστε τα παραπάνω σε ένα συμπιεσμένο αρχείο (π.χ. ZIP) και ανεβάστε το στο HELIOS.

ΟΔΗΓΙΕΣ ΓΙΑ ΤΟ KAGGLE

Πριν ξεκινήσετε τη σύνοδο σας, βεβαιωθείτε ότι έχετε επιλέξει τις παρακάτω ρυθμίσεις (Notebook options):

- ACCELERATOR: GPU P100
- LANGUAGE: Python
- PERSISTENCE: Files only
- ENVIRONMENT: Always use latest environment
- INTERNET: Internet on