



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών-Μηχανικών
Ηλεκτρονικών Υπολογιστών
Διατμηματικό Μεταπτυχιακό Πρόγραμμα
Μηχανικής Μάθησης και Επιστήμης Δεδομένων

Εξόρυξη Γνώσης από Δεδομένα

Chicago Crime Rate

Μύρων Σαμψάκης-Μπακόπουλος
Δημήτρης Λαμπράκης
Σπύρος Πούρος

Αθήνα, Φεβρουάριος 2020

Περιεχόμενα

1	Εισαγωγή	2
2	Data Augmentation	2

1 Εισαγωγή

Σκοπός της ερασίας αυτής είναι η εφαρμογή τεχνικών εξόρυξης γνώσης από το dataset Chicago Crime Rate που δίνεται από το BigQuery της Google.

Το dataset αυτό αποτελείται από δηλώσεις εγκλημάτων που γίναν στην περιοχή του Σικάγου κατά περίπου τα τελευταία 20 χρόνια. Πέραν του είδους του εγκλήματος, περιλαμβάνει πληροφορίες όπως η ώρα και η μέρα που έγινε, η περιοχή, εάν υπήρξε σύλληψη ή όχι, και διάφορα άλλα. Οι συνολικές καταχωρήσεις ανέρχονται περίπου στις 7 εκατομμύρια, καθιστώντας το αρκετά μεγάλο.

Τα εργαλεία που χρησιμοποιήθηκαν...

2 Data Augmentation