



شبكة المعلومات الجامعية  
التوثيق الإلكتروني والميكروفيلم

# بسم الله الرحمن الرحيم



**MONA MAGHRABY**



شبكة المعلومات الجامعية  
التوثيق الإلكتروني والميكروفيلم



# شبكة المعلومات الجامعية التوثيق الإلكتروني والميكروفيلم



**MONA MAGHRABY**



شبكة المعلومات الجامعية  
التوثيق الإلكتروني والميكروفيلم

# جامعة عين شمس

## التوثيق الإلكتروني والميكروفيلم

### قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها  
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات



### يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



**MONA MAGHRABY**



Cairo University

# **ENHANCING CELL-PHONES' RECEIVED SIGNAL STRENGTH PREDICTION USING DEEP LEARNING**

By

**AMR SALEH FOUAD HUSSEIN NASSAR**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communications Engineering**

FACULTY OF ENGINEERING, CAIRO UNIVERSITY  
GIZA, EGYPT  
2021

# **ENHANCING CELL-PHONES' RECEIVED SIGNAL STRENGTH PREDICTION USING DEEP LEARNING**

By  
**AMR SALEH FOUAD HUSSEIN NASSAR**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communications Engineering**

Under the Supervision of  
**Prof. Dr. Mohsen Rashwan**

.....  
Professor of Electronics and  
Communications  
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY  
GIZA, EGYPT  
2021

# **ENHANCING CELL-PHONES' RECEIVED SIGNAL STRENGTH PREDICTION USING DEEP LEARNING**

By  
**AMR SALEH FOUAD HUSSEIN NASSAR**

A Thesis Submitted to the  
Faculty of Engineering at Cairo University  
in Partial Fulfillment of the  
Requirements for the Degree of  
**MASTER OF SCIENCE**  
in  
**Electronics and Communications Engineering**

Approved by the Examining Committee

---

**Prof. Dr. Prof. Mohsen Rashwan**

**Thesis Main Advisor**

---

**Prof. Dr. Hazim Tawfik**

**Internal Examiner**

---

**Prof. Dr. Prof. Fadel Digham**

R&D director at NTRA

**External Examiner**

FACULTY OF ENGINEERING, CAIRO UNIVERSITY  
GIZA, EGYPT  
2021

**Engineer's Name:** AMR SALEH FOUAD HUSSEIN NASSAR  
**Date of Birth:** 20/01/1992  
**Nationality:** Egyptian  
**E-mail:** Eng.amrsaleh@gmail.com  
**Phone:** +201004799470  
**Address:** 177 Hay El Ashgar, 6 October City, Gizah  
**Registration Date:** 01/03/2015  
**Awarding Date:** ....../....../2021  
**Degree:** Master of Science  
**Department:** Electronics and Electric Communications Engineering



**Supervisors:**  
Prof. Mohsen Rashwan

**Examiners:**  
Prof. Mohsen Rashwan (Thesis main advisor)  
Prof. Hazim Tawfik (Internal examiner)  
Prof. Fadel Digham (External examiner)  
**R&D director at NTRA**

**Title of Thesis:**

## ENHANCING CELL-PHONES' RECEIVED SIGNAL STRENGTH PREDICTION USING DEEP LEARNING

**Key Words:**

Machine Learning; Reference Signal Strength; Telecom Optimization; Channel Model; Internet of Things

**Summary:**

While mobile operators invest in providing the best quality of service (QOS) to its customers, a live visibility on the actual QOS at the customer end is often needed. Mobile operators rely on drive test to measure QOS at user level thus identify the service level. When this visibility is inaccurate or not live, detecting and acting on customer problems can take lengthy timeframes. The thesis proposes machine learning models using huge historical dataset collected from actual filed readings to predict the QOS received at the customer level indifferent locations. Five ML approaches are examined, and the results were compared to identify the ML model that can offer higher prediction accuracy for QOS. Then Clustering ML model was built to divide the coverage area into small areas such that probe devices can be used to collect field readings from specific locations to improve the predictive model.

## **Disclaimer**

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute.

I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Amr Saleh Fouad Hussein Nassar

Date:    /    /2021

Signature:

## **Dedication**

I dedicate this thesis, with deep respect and great love, to my parents. You not only gave me the gift of life, but an unrelenting passion to live it fully. For that I am very grateful.

## **Acknowledgments**

First of all, I would like to thank God - The Provider- the one whose blessings and giving upon me is uncountable. This thesis would not have been accomplished if it was not for the help of my professor Prof. Mohsen A Rashwan, I cannot express enough thanks and gratitude for his support, encouragement, and guidance. Finally, I'd like to present a special thanks to my parents for everything.

# Table of Contents

|                                                                                |             |
|--------------------------------------------------------------------------------|-------------|
| <b>Disclaimer</b>                                                              | <b>i</b>    |
| <b>Dedication</b>                                                              | <b>ii</b>   |
| <b>Acknowledgements</b>                                                        | <b>iii</b>  |
| <b>Table of Contents</b>                                                       | <b>iv</b>   |
| <b>List of Tables</b>                                                          | <b>viii</b> |
| <b>List of Figures</b>                                                         | <b>ix</b>   |
| <b>Abstract</b>                                                                | <b>xi</b>   |
| <b>1 Introduction</b>                                                          | <b>1</b>    |
| 1.1 Problem Formulation . . . . .                                              | 1           |
| 1.2 Motivation . . . . .                                                       | 1           |
| 1.3 Deep Learning as a solution . . . . .                                      | 2           |
| 1.4 Thesis Organization . . . . .                                              | 4           |
| <b>2 Literature Survey and Theory</b>                                          | <b>6</b>    |
| 2.1 Introduction . . . . .                                                     | 6           |
| 2.2 Related work . . . . .                                                     | 6           |
| 2.2.1 Comparative study . . . . .                                              | 6           |
| 2.2.2 Estimating power from transmission distance . . . . .                    | 9           |
| 2.2.2.1 Deep Learning for RSS estimation . . . . .                             | 9           |
| 2.2.2.2 Other Models for RSS estimation . . . . .                              | 11          |
| 2.2.3 Using crowd-sourcing data . . . . .                                      | 15          |
| 2.2.4 Adding more input parameters . . . . .                                   | 16          |
| 2.2.4.1 Impact of adding meteorological parameters on RSS prediction . . . . . | 16          |
| 2.2.4.2 Impact of adding terrain profile on RSS prediction . . . . .           | 16          |
| 2.2.4.3 Impact of adding weather data on RSS prediction . . . . .              | 17          |
| 2.2.5 Predicting transmitter localization using RSS reading . . . . .          | 18          |
| 2.2.6 Summary of literature . . . . .                                          | 21          |
| 2.3 Telecommunications background . . . . .                                    | 22          |
| 2.3.1 Importance of data rates in LTE . . . . .                                | 22          |
| 2.4 Differences between 4G and 3G in QOS . . . . .                             | 23          |
| 2.4.1 The coverage difference between 4G and 3G . . . . .                      | 23          |
| 2.5 Measuring QOS . . . . .                                                    | 24          |
| 2.5.1 Down-link throughput and Up-link throughput . . . . .                    | 24          |
| 2.5.2 Received Signal Strength Indicator . . . . .                             | 24          |
| 2.5.3 Reference Signal Received Power . . . . .                                | 25          |
| 2.5.4 Reference Signal Received Quality . . . . .                              | 25          |

|          |                                               |           |
|----------|-----------------------------------------------|-----------|
| 2.5.5    | Signal to Interference and Noise Ratio (SINR) | 25        |
| 2.6      | QOS ML Prediction Model                       | 26        |
| 2.7      | Machine Learning                              | 26        |
| 2.7.1    | Machine Learning Algorithms                   | 26        |
| 2.7.1.1  | Supervised Learning                           | 26        |
| 2.7.1.2  | Unsupervised Learning                         | 34        |
| 2.7.1.3  | Reinforcement Learning                        | 36        |
| 2.7.2    | Over-fitting and Under-fitting                | 36        |
| 2.7.3    | Optimization Algorithms                       | 37        |
| 2.7.4    | Deep Learning                                 | 38        |
| 2.7.4.1  | Deep Learning vs Machine Learning             | 38        |
| 2.8      | Frameworks and Software                       | 39        |
| 2.8.1    | Programming Languages                         | 39        |
| 2.8.1.1  | Python                                        | 39        |
| 2.8.1.2  | R                                             | 39        |
| 2.8.1.3  | Java                                          | 39        |
| 2.8.2    | Frameworks and Libraries                      | 39        |
| 2.8.2.1  | TensorFlow                                    | 40        |
| 2.8.2.2  | Theano                                        | 41        |
| 2.8.2.3  | Pandas                                        | 41        |
| 2.8.2.4  | Numpy                                         | 41        |
| 2.8.2.5  | Keras                                         | 41        |
| 2.9      | Methods to tackle a Regression problem        | 41        |
| 2.9.1    | Linear and Polynomial Regression              | 41        |
| 2.9.2    | Pros of using Regression Models               | 42        |
| 2.9.3    | Cons of using Regression Models               | 42        |
| 2.9.4    | Neural Network                                | 42        |
| 2.9.5    | Pros of using Neural Networks                 | 42        |
| 2.9.6    | Cons of using Neural Networks                 | 42        |
| 2.9.7    | Regression Trees and Random Forests           | 42        |
| 2.9.8    | Pros of using Trees and Random Forests        | 43        |
| 2.9.9    | Cons of using Trees and Random Forests        | 43        |
| 2.10     | Ways of Benchmarking results                  | 43        |
| 2.10.1   | R-Squared                                     | 43        |
| 2.10.2   | RMSE                                          | 43        |
| 2.11     | ML models in practice                         | 44        |
| <b>3</b> | <b>Implementation And Results</b>             | <b>45</b> |
| 3.1      | Data-sets                                     | 45        |
| 3.1.1    | Getting the data                              | 45        |
| 3.1.2    | New Zealand's field data                      | 45        |
| 3.1.3    | Egypt's field Data                            | 45        |
| 3.2      | Data Investigation                            | 45        |
| 3.2.1    | Data-set Geo-Analysis                         | 46        |
| 3.2.1.1  | Plotting a Single Site                        | 46        |
| 3.2.1.2  | Circle of Coverage                            | 46        |
| 3.2.2    | Cell level Analysis                           | 47        |

|       |         |                                                                             |    |
|-------|---------|-----------------------------------------------------------------------------|----|
|       | 3.2.2.1 | Cell Size Estimation . . . . .                                              | 47 |
|       | 3.2.2.2 | Sample Size Per Cell . . . . .                                              | 48 |
| 3.2.3 |         | Outlier Analysis . . . . .                                                  | 48 |
|       | 3.2.3.1 | Excellent Signal At Far Distance . . . . .                                  | 48 |
|       | 3.2.3.2 | Poor Signal At Short Distances . . . . .                                    | 49 |
|       | 3.2.3.3 | Small Sized Cells . . . . .                                                 | 49 |
| 3.2.4 |         | Inspecting the Dependency between Variables . . . . .                       | 50 |
| 3.3   |         | Data Pre-processing . . . . .                                               | 50 |
|       | 3.3.1   | Data formation . . . . .                                                    | 51 |
|       | 3.3.2   | Geo-coding assistance . . . . .                                             | 51 |
|       | 3.3.3   | Staging layer . . . . .                                                     | 51 |
| 3.4   |         | Platforms and tools . . . . .                                               | 51 |
|       | 3.4.1   | Netezza Database . . . . .                                                  | 51 |
|       | 3.4.2   | Programming Language and libraries . . . . .                                | 51 |
| 3.5   |         | Data Algorithms . . . . .                                                   | 52 |
|       | 3.5.1   | Deep Learning . . . . .                                                     | 52 |
|       | 3.5.2   | Training Data Model . . . . .                                               | 52 |
|       | 3.5.3   | Customizing the DL model . . . . .                                          | 53 |
|       | 3.5.4   | Isolating training data-set . . . . .                                       | 54 |
|       | 3.5.5   | Deep Learning Model Results . . . . .                                       | 54 |
|       | 3.5.6   | Results Visualization . . . . .                                             | 54 |
| 3.6   |         | Comparing results with previous research . . . . .                          | 56 |
|       | 3.6.1   | Relation between distance and RSRP in previous research . . . . .           | 56 |
| 3.7   |         | Results Bench-marking . . . . .                                             | 57 |
|       | 3.7.1   | Theoretical Analysis . . . . .                                              | 57 |
|       |         | 3.7.1.1 Theoretical calculations . . . . .                                  | 57 |
|       |         | 3.7.1.2 Theoretical method accuracy . . . . .                               | 58 |
|       | 3.7.2   | Linear and Polynomial Regression tests . . . . .                            | 58 |
|       |         | 3.7.2.1 Test 1 - Linear Regression . . . . .                                | 58 |
|       |         | 3.7.2.2 Test 2 - Second Degree Polynomial Regression . . . . .              | 58 |
|       |         | 3.7.2.3 Test 3 - Third Degree Polynomial Regression . . . . .               | 59 |
|       |         | 3.7.2.4 Test 4 - Fourth Degree Polynomial Regression . . . . .              | 60 |
|       | 3.7.3   | Comparing DL with LR and PR . . . . .                                       | 60 |
|       | 3.7.4   | The effect of cut-over at 5Km distance . . . . .                            | 60 |
|       |         | 3.7.4.1 Test 5 - Linear Regression below 5Km . . . . .                      | 61 |
|       |         | 3.7.4.2 Test 6 - Second Degree Polynomial Regression below<br>5Km . . . . . | 61 |
|       |         | 3.7.4.3 Test 7 - Third Degree Polynomial Regression below 5Km               | 62 |
|       |         | 3.7.4.4 Test 8 - Fourth Degree Polynomial Regression below 5Km              | 63 |
|       |         | 3.7.4.5 Test 9 - Fifth Degree Polynomial Regression below 5Km               | 63 |
|       |         | 3.7.4.6 Test 10 - Deep learning model below 5Km . . . . .                   | 64 |
|       | 3.7.5   | Comparing DL with LR and PR for data-set less than 5Km . . . . .            | 64 |
| 3.8   |         | Computation comparison between the models . . . . .                         | 64 |
| 3.9   |         | Results Summary . . . . .                                                   | 65 |
| 3.10  |         | Improving the prediction through K-means clustering ML technique . . . . .  | 66 |

|          |                                                                     |           |
|----------|---------------------------------------------------------------------|-----------|
| <b>4</b> | <b>Region Clustering and probes</b>                                 | <b>67</b> |
| 4.1      | Network probes solution . . . . .                                   | 67        |
| 4.2      | Clustering . . . . .                                                | 67        |
| 4.2.1    | K-Means Clustering . . . . .                                        | 67        |
| 4.3      | Different cluster Sizes . . . . .                                   | 68        |
| 4.3.1    | Ten clusters visualization . . . . .                                | 68        |
| 4.3.2    | One Thousand Cluster Assumption . . . . .                           | 68        |
| 4.3.3    | Five Thousand Clusters Model . . . . .                              | 69        |
| 4.3.4    | Five Thousand Clusters Model Results . . . . .                      | 70        |
| 4.3.5    | Ten Thousand Clusters Model . . . . .                               | 70        |
| 4.3.6    | Real-time analysis using probes . . . . .                           | 70        |
| 4.3.7    | Extreme Clustering with 33K clusters . . . . .                      | 71        |
| 4.3.8    | Calculating Separation Distance within a cluster . . . . .          | 71        |
| 4.3.9    | Calculating Separation Distance for different clustering models . . | 71        |
| 4.3.10   | Clustering trial results . . . . .                                  | 72        |
| 4.4      | Plotting clusters trials on a curve . . . . .                       | 72        |
| 4.5      | Comparing Clustering with previous results . . . . .                | 72        |
| 4.6      | Use-cases and Future Work . . . . .                                 | 74        |
| <b>5</b> | <b>Conclusion, Use-Cases And Future Work</b>                        | <b>75</b> |
| 5.1      | Conclusion . . . . .                                                | 75        |
| 5.2      | Commercial use-cases . . . . .                                      | 75        |
| 5.2.1    | Network Visibility Use-Cases . . . . .                              | 75        |
| 5.2.1.1  | Single Customer Analysis . . . . .                                  | 76        |
| 5.2.1.2  | Geo-location Analysis with probes . . . . .                         | 76        |
| 5.2.1.3  | Direct Selling . . . . .                                            | 76        |
| 5.2.1.4  | Regulatory Monitoring . . . . .                                     | 76        |
| 5.2.1.5  | Improve Network Planning . . . . .                                  | 76        |
| 5.2.2    | Estimated Revenue from New Sites . . . . .                          | 76        |
| 5.2.2.1  | Retain Existing Customers . . . . .                                 | 76        |
| 5.2.2.2  | Better Coverage at Blind Spots . . . . .                            | 77        |
| 5.2.2.3  | Minimize Churn . . . . .                                            | 77        |
| 5.2.2.4  | Get Rid of Costly Customers . . . . .                               | 77        |
| 5.3      | Future Work . . . . .                                               | 78        |
| 5.3.1    | Add Input Parameters . . . . .                                      | 78        |
| 5.3.1.1  | Weather Condition . . . . .                                         | 78        |
| 5.3.1.2  | Geography . . . . .                                                 | 78        |
| 5.3.1.3  | Number of Active Users . . . . .                                    | 78        |
| 5.3.1.4  | BS Configurations . . . . .                                         | 78        |
| 5.3.2    | Use Bigger Data-set . . . . .                                       | 79        |
| 5.3.3    | Examine Other Attributes . . . . .                                  | 79        |
|          | <b>Glossary</b>                                                     | <b>80</b> |
|          | <b>References</b>                                                   | <b>81</b> |

# List of Tables

|     |                                                                       |    |
|-----|-----------------------------------------------------------------------|----|
| 3.1 | Comparing DL with LR and PR results . . . . .                         | 60 |
| 3.2 | DL, LR and PR Results comparison in 5KM range . . . . .               | 64 |
| 4.1 | Average Separation Distance for different Clustering models . . . . . | 71 |
| 4.2 | Benchmarking Clustering results with other ML models . . . . .        | 73 |