RESPIRATORY SOCIETY CONGRESS – SARAJEVO 2013

Faculty of Electrical Engineering
Department of Computer Science

Faculty of Medicine
Center for Genetics

COMPUTER AIDED LUNG CANCER CLASSIFICATION OF MUTATED EGFR EXONS USING ARTIFICIAL INTELLIGENCE METHODS


Prof. Zikrija Avdagic, DSc[1]


PhD Aida Saracevic[2]


MSc Dino Keco[3]


MA Aja Avdagic[4]


MSc Dusanka Boskovic[5]

[1,3,5]Faculty of Electrical Engineering-University of Sarajevo, Bosnia and Herzegovina
[2]Center for Genetics, Medical Faculty-University of Sarajevo, Bosnia and Herzegovina
[4]Faculty of Medicine-Ludwig Maximilian University of Munich, Germany

Problem Domain Description

2013-2014

Research fields

Artificial Intelligence, Bioinformatics, Molecular Genetics, Biomedical Engineering

Molecular diagnostics is a rapidly advancing field in which insights into:

microarray gene expression analysis,
microarray miRNA expression analysis,
cell signaling pathways,

may provide invaluable information on:

disease pathology,
disease progression,
resistance to treatment,
response to cellular microenvironments,

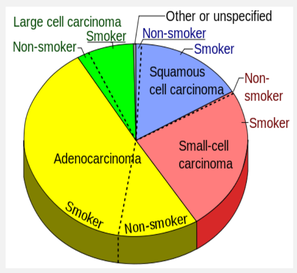
and ultimately may need fusion of:

artificial intelligent methods,
bioinformatics,
mathematics & statistics

TO IMPROVE EARLY DIAGNOSIS AND CREATE INNOVATIVE THERAPEUTIC APPROACHES FOR CANCER.

This research has as its goal to improve diagnosis of Non-Small Cell Lung Cancer(NSCLC) based on sample patients' data with microdeletion mutations extracted from online EGFR mutation database, and samples data with microdelition mutations generated in own generator reqiured for simulation.We have developed an integrated software suit based on module for preprocessing data (extraction, encoding, and normalization), module for exon microdelitions generation (statistical data base), module for training/learning of artificial neural networks (ANN), and module for postprocessing (classification, and evaluation). We have made experiments on eleven different training/learnig algorithms in combination with different number of cells, layers, and activation functions. The best results have been achieved with cascade-forward backpropagation algorithm based on Levenberg-Marquardt learning mechanism, including best performance (error 9.3083e-031) with the minimum epochs (training iterations 8), and the regression fit curves (trainig,validation and testing R=1). Dividing trainig sets into smaller subsets using ensemble-voting methods on each exons would improve system reliabilty in situations of possible single neural network fail.

Lung Cancer



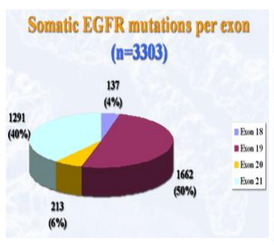
The main types of lung cancer are:
1. non-small-cell lung carcinoma (NSCLC), and
2. small-cell lung carcinoma (SCLC), also called oat cell cancer.

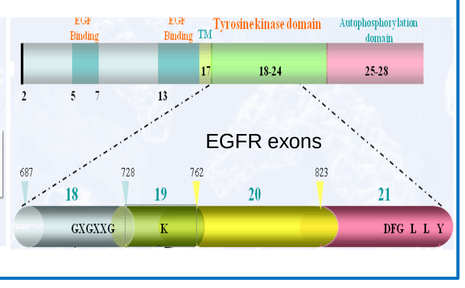
There are three common forms of NSCLC:
1.1 Adenocarcinomas are often found in an outer area of the lung.
1.2 Squamous cell carcinomas are usually found in the center of the lung next to an air tube (bronchus).
1.3 Large cell carcinomas can occur in any part of the lung. They tend to grow and spread faster than the other two types.

There are two different types of SCLC:
2.1 Small cell carcinoma (oat cell cancer).
2.2 Combined small cell carcinoma.

Statistics on Lung Cancer

- The most common cause of lung cancer is long-term exposure to tobacco smoke which causes 80%-90% of lung cancers.
- Nonsmokers account for 10-15% of lung cancer cases and these cases are often attributed to a combination of genetic factors, radon gas, asbestos and air pollution including second hand smoke.
- Different combinations of mutations (micro-deletions) exist within the EGFR kinase, and the most frequently observed mutations are on the exon 18, 19,20 and 21.



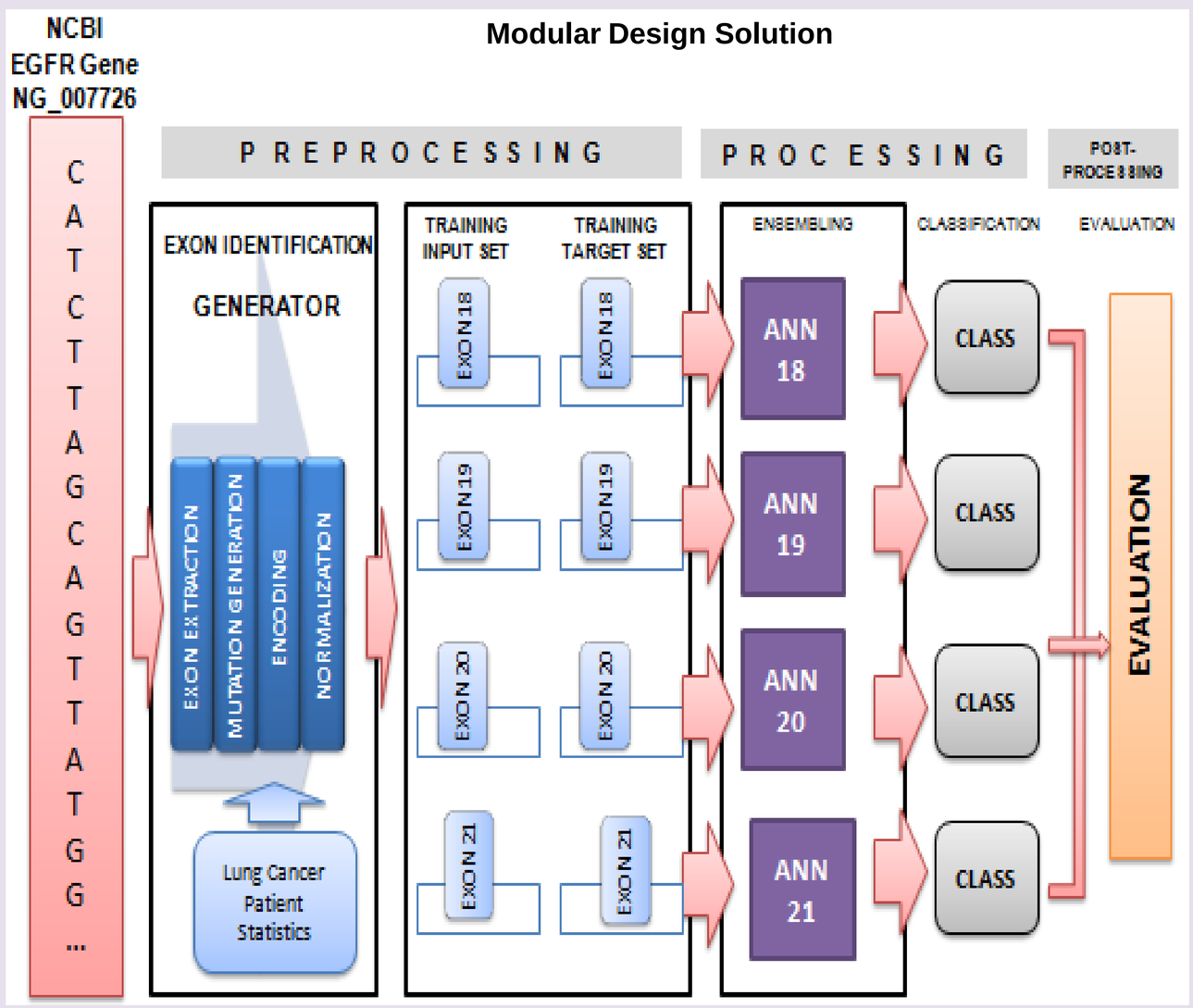


Statistics for Mutation Generator

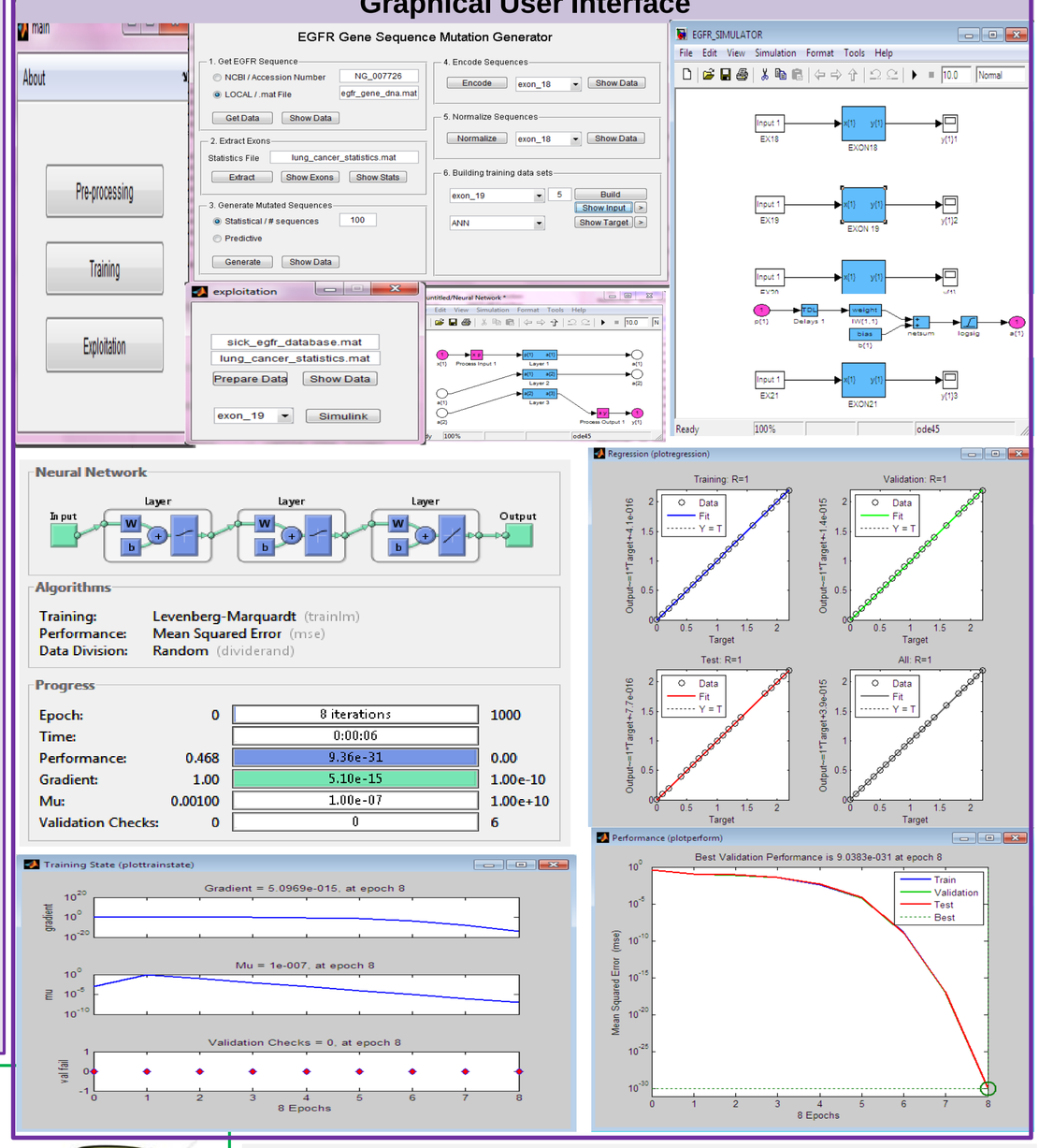
lung_cancer_mutation_statistics — POSITIONS OF MUTATIONS | PATIENTS | EXONS | CONVERSIONS

	1	2	3	4	5	6	7	8
1	1	160691	160789	50	64	166	19	'-'
2	2	160691	160789	51	65	60	19	'-'
3	3	160691	160789	69	92	1	19	'-'
4	4	160691	160789	55	72	33	19	'-'
5	5	160691	160789	55	69	6	19	'-'
6	6	160691	160789	54	71	7	19	'-'
7	7	160691	160789	52	66	7	19	'-'
8	8	160691	160789	53	67	2	19	'-'
9	9	160691	160789	53	70	2	19	'-'
10	10	160691	160789	52	69	3	19	'-'
11	11	160691	160789	54	62	2	19	'-'
12	12	160691	160789	54	68	1	19	'-'
13	13	160691	160789	60	68	1	19	'-'
14	14	160691	160789	68	91	2	19	'-'
15	15	167262	167447	25	26	2	20	'-'
16	16	160691	160789	51	56	1	19	'-'
17	17	159890	160012	96	97	1	18	'-'
18	18	160691	160789	53	62	1	19	'-'
19	19	160691	160789	69	70	1	19	'-'
20	20	160691	160789	50	51	2	19	'-'
21	21	160691	160789	55	66	3	19	'-'
22	22	160691	160789	44	51	1	19	'-'
23	23	177688	177843	103	103	305	21	'G>T'
24	24	177688	177843	112	112	7	21	'A>T'
25	25	177688	177843	7	7	2	21	'G>A'
26	26	177688	177843	103	104	2	21	'TG>GT'

Modular Design Solution



Graphical User Interface



Epidermal Growth Factor Receptor (EGFR) Targeted Therapy

New cancer treatments target and turn off EGFR signals. These therapies use antibodies that recognize only EGFR, stick to it, and block it from sending messages. By interrupting the signals, cancer cells are no longer told to overgrow, and eventually die. These treatments are called EGFR-targeted therapies.

•New cancer treatments such as receptor tyrosine kinase inhibitors(RTKIs) that target and turn off EGFR signals.
•By inhibiting the ATP, formation of phosphotyrosine residues inEGFR is not possible and the signal cascades are not initiated.
•Gefitinib (Iressa) inhibits EGFR tyrosine kinase domain by binding to the adenosine triphosphate(ATP)-binding site of the enzyme. Thus the function of the EGFR tyrosine kinase in activating the anti-apoptotic RAS signal transduction cascade is inhibited, and malignant cells are inhibited.
•Erloinib(Tarceva) specifically targets the EGFR tyrosine kinase, which is highly expressed and occasionally mutated. It binds in a reversible fashion to the adenosine triphosphate(ATP) binding site of the receptor.

Patient ID	Pat N0.	Mutation. Type	Nucleotide change	Researcher Author	Gender	Age	Ethnicity/ Putative	Smoking status	Treatment
2051	19	Microdeletion	c.2235_2249del	Sasaki H, Endo K, et al.	F		Putative Japanese		
2148	19	Microdeletion	c.2235_2249del	Takano T, Ohe Y, et al.	F	70	Putative Japanese	never	Gefitinib
2149	19	Microdeletion	c.2235_2249del	Takano T, Ohe Y, et al.	M	53	Putative Japanese	former	Gefitinib
2150	19	Microdeletion	c.2235_2249del	Takano T, Ohe Y, et al.	F	57	Putative Japanese	never	Gefitinib
2151	19	Microdeletion	c.2235_2249del	Takano T, Ohe Y, et al.	M	70	Putative Japanese	never	Gefitinib
2132	19	Microdeletion	c.2235_2249del	Eberhard DA, Johnson BE, et al.	M		N/A	Former	Erlotinib+Chemotherapy
2127	19	Microdeletion	c.2235_2249del	Eberhard DA, Johnson BE, et al.	F		N/A	Former	Erlotinib+Chemotherapy
2128	19	Microdeletion	c.2235_2249del	Eberhard DA, Johnson BE, et al.	F		N/A	Former	Erlotinib+Chemotherapy
2129	19	Microdeletion	c.2235_2249del	Eberhard DA, Johnson BE, et al.	M		N/A	Former	Erlotinib+Chemotherapy

Future Plan

In order to complete this work our next task is to develop a more powerful mutation prediction generator for exon 18, 19, and 20 covering microdelition mutations that take place over the nucleotides in consecutive order (the sum of all mutations with one deletion, two deletions and so on to the number of microdelition corresponding to the length exon).
Mutations on exon 21 will be based on translations of one kind nucleotide to another (G -> T, G->A, A -> T), and translation of a pair of nucleotides to second pair of nucleotides (TG -> GT).This work is only first step within our:

LONG TERM RESEARCH

GLOBAL APPROACH
based on
microARRRAY, microRNA AND CELL SIGNALLING PATHWAYS
using
FUSION OF ARTIFICIAL INTELLIGENCE METHODS
to improve
EARLY DIAGNOSIS AND INNOVATIVE METHODS FOR CANCER

References

EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.
Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H ...
Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.
Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, ...
Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways.
Sordella R, Bell DW, Haber DA, Settleman J.
EGFR SIGNALING NETWORKS IN CANCER THERAPY
John D. Haley
Receptor Tyrosine Kinases and Targeted Cancer Therapeutics
Kenji TAKEUCHI* and Fumiaki ITO
Ensmbling of EGFR Mutations'Based on Artificial Neural Networks for Improved Diagnosis of Non-Small Cell Lung
Emanuel Adetiba, Frank A.Ibikuline
Vještačka inteligencija&Fuzzy-Neuro-Genetika
Zikrija Avdagic
MATLAB, SIMULINK, NEURAL NETWORK TOOLBOX,BIOINFORMATIC TOOLBOX, Statistics Toolbox
Mathworks Inch