

საქართველოს ტექნიკური უნივერსიტეტი

ე.ყუბანეიშვილი, მ.მესხია

**ნეიროქსელური ტექნოლოგიები
STATISTICA სისტემის ბაზაზე**

პრაქტიკული

თბილისი 2016

შესავალი

ნეირონული ქსელები ფართოდ გამოიყენება ყველგან, სადაც საჭიროა პროგნოზირების, კლასიფიკაციის, კლასტერიზაციის, ოპტიმიზაციის, წარმოების პროცესების მართვის და სხვა პრაქტიკული ამოცანების გადაწყვეტა. ნეიროქსელური ტექნოლოგიები შეიძლება გამოყენებული იყოს დამოუკიდებლად ან მონაცემთა ანალიზის ტრადიციულ მეთოდებთან ერთად, სადაც ის წარმოადგენს მნიშვნელოვან დანამატს.

კლასიკური სტატისტიკული მეთოდების უმეტესობა ძირითადად დაკავშირებულია მათემატიკური მოდელების აგებაში, რომლებიც დაფუძნებული არიან ამა თუ იმ დაშვებებზე და თეორიულ მოსაზრებებზე. (მაგალითად, საძებნი დამოკიდებულება წრფივია, ცვლადებს გააჩნიათ ნორმალური განაწილება და ბევრი სხვა).

ნეიროქსელური მიდგომები თავისუფალია მოდელური შეზღუდვებისაგან. ის ერთნაირად გამოიყენება როგორც წრფივი, ასევე რთული არაწრფივი დამოკიდებულებების დროს. განსაკუთრებით საინტერესოა ნეიროქსელური ტექნოლოგიების გამოყენება, როცა უცნობია პარამეტრებს შორის დამოკიდებულება. გარდა ამისა, ნეიროქსელებს გააჩნიათ მეტად უნიკალური თვისება, რომელიც წარმოადგენს თვითსწავლებას. სწავლების პროცედურა შედგება სინაფსური წონების რეგულირებაში, რათა მიღწეული იყოს ცდომილების ფუნქციის მინიმიზაცია.

ნეიროქსელური პროგრამული პროდუქტი *STATISTICA Neural Networks (SNN)* წარმოადგენს ნეირონული ქსელების მოდელების შექმნის და ანალიზის მძლავრ და სწრაფ საშუალებას. ნეიროქსელის აგებისათვის საჭიროა შეირჩეს ცვლადები (რიცხვითი ან ნომინალური), რომლებიც მოქმედებენ შედეგზე. პროგრამაში გათვალისწინებულია შემავალი და გამომავალი ცვლადები. ცვლადები არის ორი სახის: დამოკიდებული (გამომავალი ცვლადები) და დამოუკიდებელი (შემავალი ცვლადები), რომლებიც გამოიყენებიან ქსელის შესასწავლად.

არსებობს აგრეთვე, კატეგორიული ცვლადები, რომლებიც შეიცავენ მთელ რიცხვებს ან ტექსტურ მნიშვნელობებს და მიეკუთვნებიან დამაჯგუფებელ ცვლადების კატეგორიას. კატეგორიული ცვლადის ტიპურ მაგალითს წარმოადგენს სქესი ორი მნიშვნელობით (მამრობითი და დედრობითი).

პროგრამაში შეიძლება გამოვიყენოთ ამოცანის ამოხსნის ორი ინსტრუმენტი: გადაწყვეტის ოსტატი (*Intelligent Problem Solver*) და ქსელის კონსტრუქტორი (*Custom Network Designet*). ამოცანის ამოსახსნელად გადაწყვეტის ოსტატს სჭირდება დიდი დრო (რომელსაც ის ამას კარგად იყენებს), განსაკუთრებით განზომილების შემცირების დროს. ქსელის კონსტრუქტორის მთავარი თვისება ის არის, რომ მისი საშუალებით შესაძლებელია ნეირონული ქსელის აგება წინასწარ მოცემული პარამეტრებით: ქსელის ტიპი (ტოპოლოგია, არქიტექტურა), ფარული შრეების რაოდენობა და სხვა.

ნეიროქსელის აგებაში მონაწილე ამონარჩევი (ობიექტები, სახეები, მოვლენები) ორი სახისაა: სასწავლო და ტესტური (საკონტროლო). სასწავლო ამონარჩევის დაკვირვებები გამოიყენებიან ქსელის სწავლებისათვის. ტესტირების ამონარჩევის დაკვირვებები არ გამოიყენებიან ქსელის სწავლების პროცესში, ისინი გამოიყენებიან სწავლების შემდგომ ქსელის საბოლოო შეფასებისათვის.

სწავლების პროცესის ავტომატური დამთავრება ხდება მაშინ, როცა ქსელი ყველა მაგალითს დაამუშავებს მოცემული სწავლების ცდომილების გათვალისწინებით. ამის შემდეგ შესაძლებელია მოცემული ნეირონული ქსელის გამოყენება ნებისმიერი სახის ახალი შემავალი მონაცემების დამუშავებისათვის.

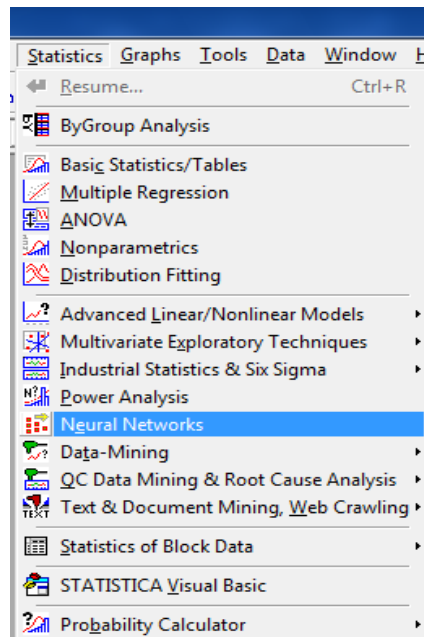
პრაქტიკული სამუშაო 1

თანამედროვე ნეიროქსელური პროგრამული პროდუქტის STATISTICA Neural Networks (SNN) ბაცნობა რეგრესიული ანალიზის მაბალითზე.

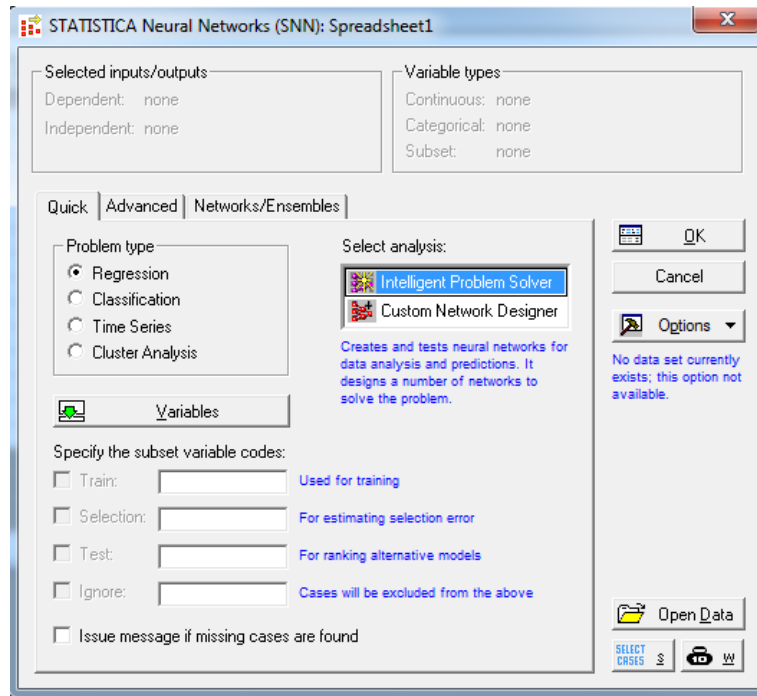
პროგრამასთან მუშაობის პროცედურა ზოგადად შედგება შემდეგი ბიჯებისაგან

ბიჯი 1. *Neural Networks*სა სტარტო პროგრამის გაცნობა.

STATISTICA Neural Networks პროგრამის გამოსაძახებლად ჩავტვირთოდ *STATISTICA 7* სისტემა და *Statistics* მენიუში მოვნიშნოთ ბრძანება *Neural Networks*.



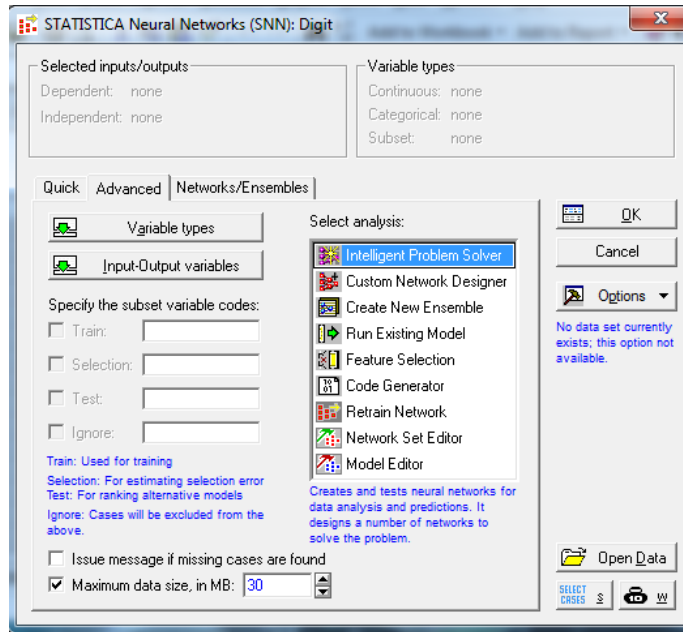
ეკრანზე გამოდის *STATISTICA Neural Networks(SNN)* პროგრამის სასტარტო ფანჯარა, რომელსაც აქვს შემდეგი სახე:



ფანჯრის **Problem type** ჩარჩოში შეგვიძლია შევარჩიოთ ანალიზის სხვადასხვა ტიპი: **Regression** (რეგრესია), **Classification** (კლასიფიკაცია), **Time Series** (დროითი მწკრივები) და **Cluster Analysis** (კლასტერული ანალიზი). შევარჩიოთ რომელიმე მათგანი, მაგალითად რეგრესია.

Select analysis ველში შევარჩიოთ ანალიზის ინსტრუმენტი **Intelligent Problem Solver** (გადაწყვეტის ოსტატი) ან **Custom Network Designer** (ქსელის კონსტრუქტორი). დამწყები მომხმარებლისათვის რეკომენდირებულია შეირჩეს გადაწყვეტის ოსტატი, ხოლო გამოცდილი მომხმარებლისათვის – ქსელის კონსტრუქტორი. შევარჩიოთ გადაწყვეტის ოსტატი.

Advanced (დამატებითი) ჩანართის ჩართვის შემდეგ



შესაძლებელია დამატებითი ინსტრუმენტების გამოყენება, მაგალითად, განზომილების შემცირება, კოდის გენერატორი და სხვა.

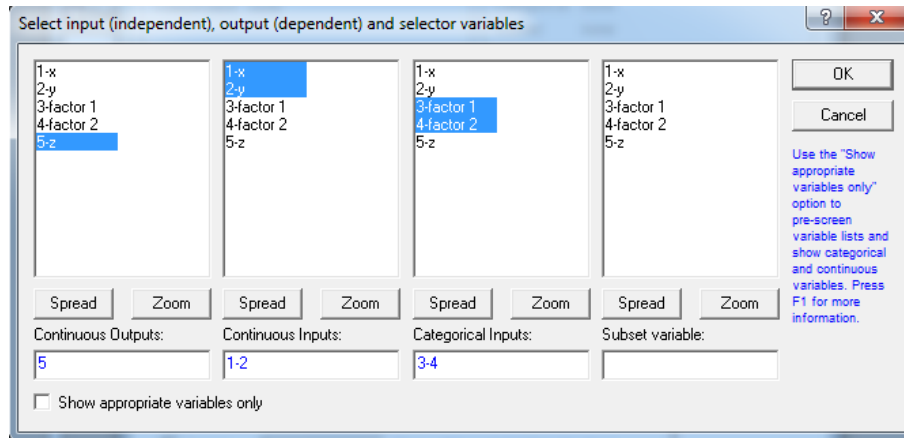
ბიჯი 2 მონაცემების ფაილის ჩატვირთვა.

თუ მონაცემთა ფაილი გახსნილი არ არის, მაშინ **STATISTICA 7** სისტემის ფანჯარაში გაეხსნათ მენიუ **File + Open Examples**. ეკრანზე გამოსულ **Open STATISTICA Data File** ფანჯარაში გაეხსნათ **Datasets** საქაღალდე და შესაბამისი ფაილის დასახელებაზე დაეაწკაპუნოთ (**Click**).

	1 x	2 y	3 factor 1	4 factor 2	5 z
1	33,75958	40,2234	m	i	1357,921
2	67,40318	53,43791	m	i	3601,639
3	84,7438	4,086428	m	d	-0,49068
4	95,52904	53,16935	s	i	148,8916
5	84,3379	14,12091	s	i	98,31806
6	33,91217	2,569658	m	i	87,4407
7	2,444533	14,12091	s	d	0,061378
8	90,99399	70,47945	m	d	0,076019
9	83,3491	10,57772	s	i	93,15478
10	71,79174	55,54979	s	d	-0,08583
11	93,82306	72,1305	m	d	-0,39306
12	97,56157	86,84347	m	i	8472,709
13	58,64742	21,05167	m	d	0,073404
14	49,04935	76,18641	m	d	0,234397

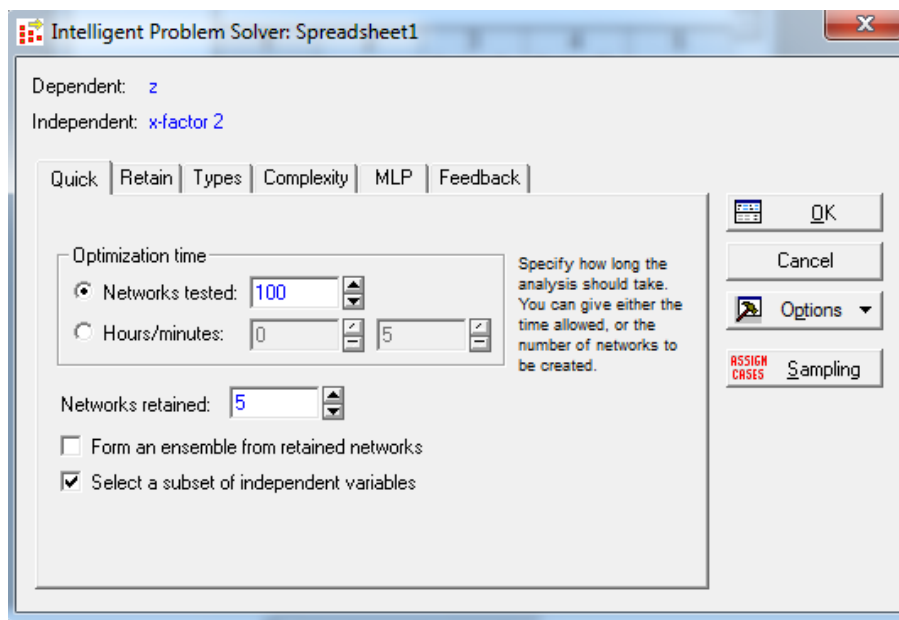
ბიჯი 3. ცვლადების მონიშვნა.

ამისათვის პროგრამის სასტარტო ფანჯარაში *Click Variables* დილაკზე და ეკრანზე გამოსულ *Select input(independent), outp(dependent) and Selector vaiiables* მონიშნოთ შემავალი (დამოუკიდებელი), გამომავალი (დამოკიდებული) და დამაჯგუფებელი ცვლადები. *Continuous outputs* (უწყვეტი გამომავალი) ველში ჩვენ შემთხვევაში, უნდა მონიშნოთ *z* ცვლადი. *Continuous inputs* (უწყვეტი შემავალი) ველში – *x* და *y* ცვლადები. *Categorical inputs* (შემავალი გატეგორიული) ველში, ჩვენი ამოცანიდან გამომდინარე, *Factor 1* და *Factor 2*. *Subset variable* (ქვესიმრავლეებად დაყოფა) არ წარმოადგენს აუცილებელს. იგი განკუთვნილია იმ ცვლადების მოსანიშნად, რომლებიც შეიცავენ კოდებს მონაცემების დასაყოფად სასწავლო და ტესტურ სიმრავლეებად. ანალიზის გასაგრძელებლად *Click OK* დილაკზე.



ბიჯი 4. ანალიზის ხანგრძლივობის და ქსელის რაოდენობის დაყენება.

ვბრუნდებით სასტარტო ფანჯარაში და კვლავ *OK*. ეკრანზე გამოდის *Intelligent Problem Solver* ფანჯარა,

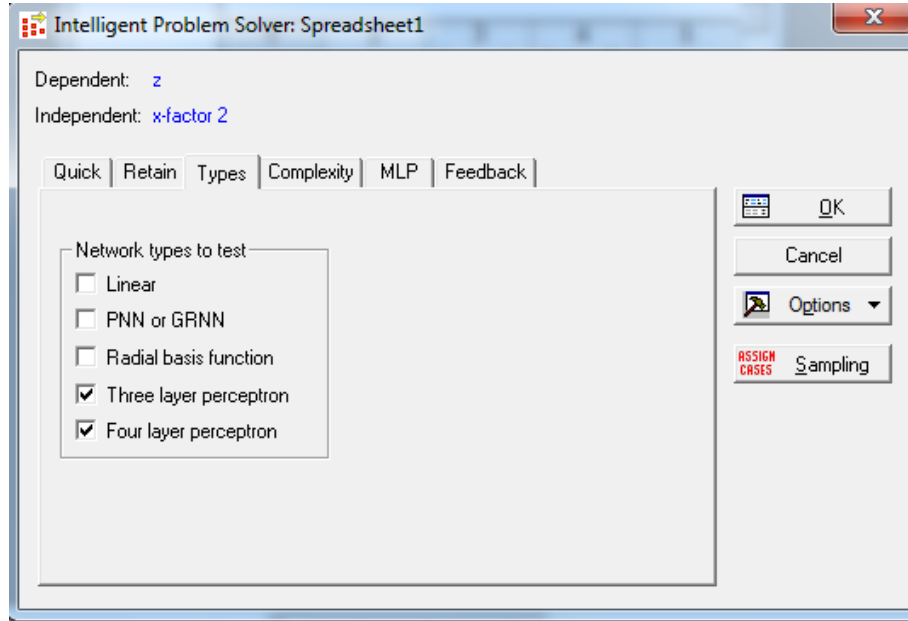


სადაც ჩავრთოდ *Quick* ჩანართი. *Networks tested* და *Networks retained* ველებში ანალიზის ტიპიდან და ამოცანიდან გამომდინარე ჩავწეროთ შესაბამისი მნიშვნელობები (100 და 5).

ბიჯი 5. ქსელის ტიპის შერჩევა.

ჩავრთოდ *Types* ჩანართი და *Network types to test* ველში მოვნიშნოთ ქსელის ან ქსელების ტიპები: *linear*(წრფივი), *PNN or GRNN* (ალბათური ნეიროქსელი ან

განზოგადოებული რეგრესიული ნეიროქსელი), **Radial basis function** (რადიალურ-ბაზისური ფუნქცია), **Three layer perceptron** (სამშრიანი რეგრესია), **Four layer perceptron** (ოთხშრიანი რეგრესია). შეძლევ **OK**.



კლასიფიკაციის ამოცანებში ძირითადად გამოიყენებიან მრავალშრიანი პერსეპტრონები, რადიალურ-ბაზისური ფუნქცია, ალბათური ნეიროქსელი (**PNN**) და წრფივი ქსელი. ამასთან შემაჯავალი მონაცემები უნდა იყოს რიცხვითი და არა ნომინალური.

რეგრესიის ამოცანებში ძირითადად გამოიყენებიან მრავალშრიანი პერსეპტრონები, რადიალურ-ბაზისური ფუნქცია, განზოგადოებული რეგრესიული ნეიროქსელი (**GRNN**) და წრფივი ქსელი.

დროითი მწკრივების ანალიზის ძირითადი მიზანია მომავალი მნიშვნელობების პროგნოზირება. როგორც წესი, პროგნოზირებადი ცვლადი რიცხვითი მნიშვნელობისაა, ამიტომ დროითი მწკრივის პროგნოზირება წარმოადგენს რეგრესიის კერძო შემთხვევას. მაგრამ **STATISTICA 7** სისტემაში ეს შეზღუდვა არ არის გათვალისწინებული და ამიტომ შესაძლებელია ნომინალური (ანუ საკლასიფიკაციური) დროითი მწკრივის პროგნოზირებაც. აქედან გამომდინარე, დროითი მწკრივის პროგნოზირების ამოცანაში გამოიყენება ყველა ტიპის ნეიროქსელი (ამოცანიდან გამომდინარე ქსელის ტიპი უნდა მიესადაგებოდეს რეგრესიას ან კლასიფიკაციას).

კლასტერიზაციის ამოცანებში ძირითადად გამოიყენებიან კოჰონენის ქსელი და მრავალშრიანი ნეირონული ქსელები. ქსელის ტიპის მონიშვნის შეძლევ **OK**.

ბიჯი 6. ქსელის სწავლების პროცედურა.

ქსელის ტიპის შერჩევის შემდეგ ეკრანზე ჯერ გამოდის სწავლების პროცედურის ფანჯარა და შემდეგ შედეგების **Results (Run Models)** ფანჯარა,

Index	Profile	Train Perf.	Select Perf.	Test Perf.	Train Error	S
1	MLP 3:3-5-1:1	0.050195	0.892404	3.567189	0.017138	0
2	MLP 3:3-8-5-1:1	0.045560	0.743112	3.572800	0.015502	0
3	MLP 4:4-1-1-1:1	0.039634	0.469559	2.258964	0.014667	0
4	MLP 4:4-3-1:1	0.084469	0.566623	1.410926	0.028697	0
5	MLP 4:4-10-4-1:1	0.116080	0.336631	3.633354	0.040126	0

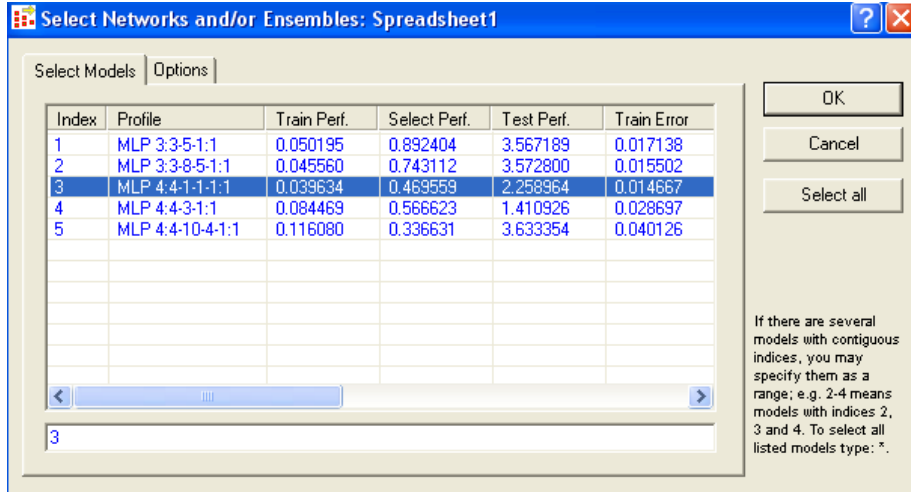
სადაც პროგრამას გამოაქვს საუკეთესო ქსელები, რომელთა მნიშვნელოვან მაჩვენებლებს წარმოადგენენ: **Profile**(ქსელის ტიპი), **TrainPerformance**(სწავლების ეფექტიანობა), **Select. Perf.** (შერჩევის ეფექტიანობა), **TestPerf.** (ტესტირების ეფექტიანობა), **Train Error**(სწავლების ცდომულება), **SelectError** (შერჩევის ცდომილება), **TestError** (ტესტირების ცდომილება).

Profile სვეტში წარმოდგენილია: ქსელის ტიპი, რომელიც აღწერილია ნეიროქსელის ტოპოლოგიით ანუ ქსელის კლასით, რომელსაც ის მიეკუთვნება (მაგალითად, **MLP**–პერსეპტრონი, შემავალი და გამომავალი ცვლადების რაოდენობა, ფარული შრეების რაოდენობა და თითოეული შრის ელემენტების რაოდენობა).

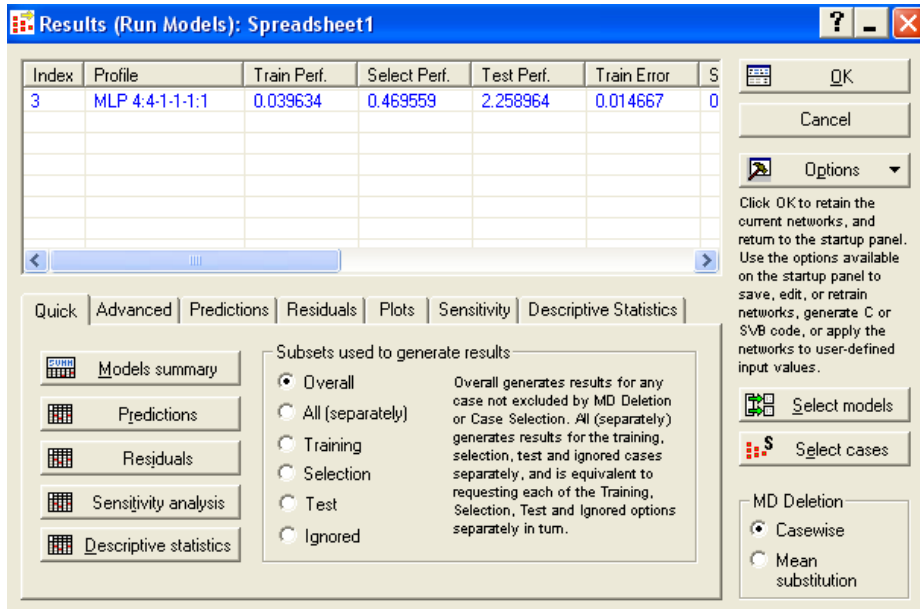
მიღებულ შედეგებიდან ჩანს, რომ ქსელის ფარული შრის ელემენტების რაოდენობის გაზრდით, იზრდება შედეგის სიზუსტე. უნდა აღინიშნოს, რომ ყველა მოდელს აქვს შესაძლებლობა ყოველ ახალ დაკვირვებაზე განახოვადოს შედეგი.

ბიჯი 7. ნეირონული ქსელის შერჩევა.

Click **Select models** (მოდელის მონიშვნა) ღილაკზე და ეკრანზე გამოსულ ცხრილში მოვნიშნოთ მე-3 ქსელი, რადგან მას გააჩნია ყველაზე კარგი შედეგი.



შემდეგ *Click OK.*



შედეგების **Results(Run Models)** ფანჯარის **Quick** ჩანართში *Click Descriptive statistics*(აღწერითი სტატისტიკა)-ის ღილაკზე. ეკრანზე გამოდის რეგრესიის შეცდომების ცხრილი,

		Regression (3) (Spreadsheet1)		
		z.3		
Data Mean	989,968			
Data S.D.	2282,389			
Error Mean	-4,146			
Error S.D.	201,133			
Abs E. Mean	155,630			
S.D. Ratio	0,088			
Correlation	0,996			

სადაც წარმოდგენილია **Data Mean** (მონაცემების საშუალო სიდიდე), **Data S.D.** (მონაცემების საშუალო კვადრატული გადახრა), **Error Mean** (პროგნოზის საშუალო ცდომილება), **Error S.D.** (პროგნოზის ცდომილების საშუალო კვადრატული გადახრა), **Abs. E. Mean** (პროგნოზის საშუალო აბსოლუტური ცდომილება), **S.D.Ratio** (პროგნოზის საშუალო კვადრატული გადახრის ცდომილების ფარდობა მონაცემების საშუალო კვადრატულ გადახრასთან), **Correlation** (კორელაციის მნიშვნელობა რეალურ და პროგნოზირებად მნიშვნელობებს შორის).

S.D.Ratio წარმოადგენს ერთ-ერთ ყველაზე უფრო მნიშვნელოვან და გამოყენებად მაჩვენებელს. მისი მნიშვნელობა $[0;1]$ დიაპაზონში იცვლება და დამოკიდებულია ყოველი მნიშვნელობის ცდომილებაზე (ნიშანზე არ არის დამოკიდებული). ამასთან ის საკმაოდ მკმნობიარეა ამოვარდნების მიმართ.

ემპირიული გამოკვლევების შედეგად ჩამოყალიბდა გარკვეული წესი, რომლის თანახმად თუ **S.D.Ratio** $< 0,2$ სიდიდეზე, მაშინ ითვლება, რომ ნეიროქსელი შერჩეულია კარგად. თუ **S.D.Ratio** მნიშვნელობა მოთავსებულია $[0,2;0,4]$ დიაპაზონში, მაშინ ითვლება, რომ ქსელი შერჩეულია დამაკმაყოფილებლად, ხოლო როცა **S.D.Ratio** $> 0,4$ სიდიდეზე, მაშინ ქსელი ცუდად არის შერჩეული.

მიღებულ შედეგებიდან ჩანს, რომ მე-3 ქსელის სტანდარტული გადახრა (**S.D.Ratio**) 0,088 ტოლია. ე.ი. ქსელი კარგად არის შერჩეული.

ნეიროქსელის შერჩევის ერთ-ერთ მნიშვნელოვან მეთოდს წარმოადგენს **კროს-შემოწმების** მეთოდი, სადაც გამოიყენება ტესტური ამონარჩევი. ტესტური ამონარჩევით მიღებული შედეგის სიზუსტე დარდება სასწავლო ამონარჩევით მიღებული შედეგის სიზუსტეს. თუ ტესტური ამონარჩევი იძლევა დაახლოებით იგივე შედეგს, რასაც სასწავლო ამონარჩევი, მაშინ ითვლება, რომ მოდელმა გაიარა კროს-შემოწმება.

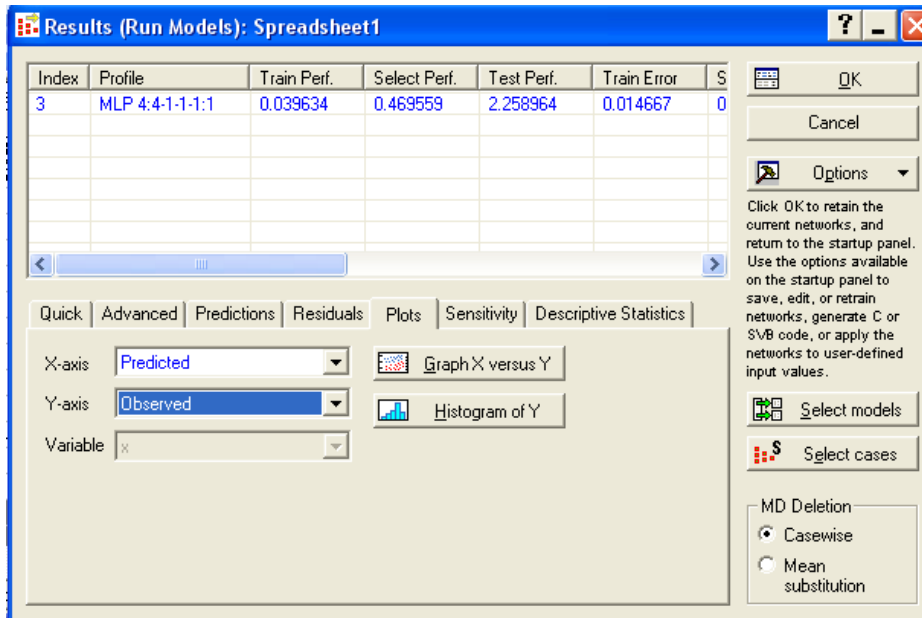
თუ შედეგების **Results(Run Models)** ფანჯრის **Subets used to generale resultsc** (შერჩევის შედეგების გამოყვანა)-ის ველში **Overall** (სრული) ოპციის მაგივრად ჩავრთავთ **All (separately)** (ყველა ცალკე) ოპციას, მაშინ ეს მოგვცემს საშუალებას მივიღოთ ცხრილი, სადაც შერჩევის შედეგები ყველა მოდელისათვის ცალკე არის წარმოდგენილი.

	Regression (1-5) (Spreadsheet1)				
	z.1	z.2	z.3	z.4	z.5
Data Mean	989,968	989,968	989,968	989,968	989,968
Data S.D.	2282,389	2282,389	2282,389	2282,389	2282,389
Error Mean	-21,598	26,533	-4,146	-78,328	-32,846
Error S.D.	285,010	256,302	201,133	280,211	295,892
Abs E. Mean	201,060	166,186	155,630	254,778	242,745
S.D. Ratio	0,125	0,112	0,088	0,123	0,130
Correlation	0,992	0,994	0,996	0,992	0,992

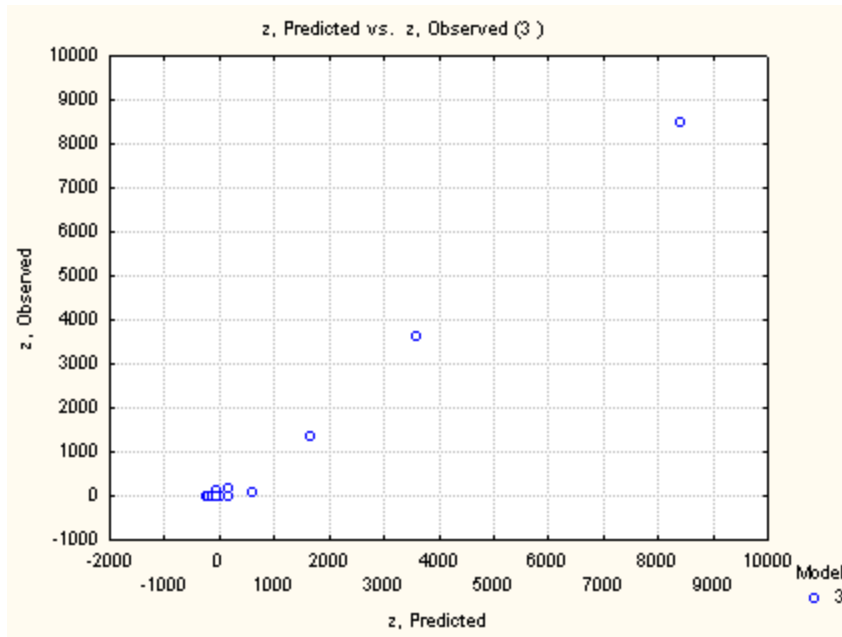
როგორც ცხრილიდან ჩანს **S.D.Ratio** სიდიდე ყველა ამ მოდელისათვის თითქმის ერთნაირია და ნაკლებია 0,2 სიდიდეზე.

საწყისი მონაცემების დაყოფა სასწავლო და ტესტურ სიმრავლეებად ხორციელდება მონაცემების დაყოფით გარკვეული პროპორციით. მაგალითად, სასწავლო ამონარჩევი – 2/3, ხოლო ტესტური – 1/3. გაჩუმების პრინციპით დაკვირვებები იყოფა სამ ქვესიმრავლედ 2 : 1 : 1 შეფარდებით. ასე მაგალითად, თუ გვაქვს 50 სასწავლო დაკვირვება, მაშინ 25 იქნება საკონტროლო და შემდეგი 25 – სატესტო.

დაკვირვების გრაფიკის ასაგებად ჩავრთოდ ჩანართ **Plots** (გრაფიკი) . და **X-axis** (დერძზე)-ის ველში შევარჩიეთ **Predicted**(პროგნოზირება), ხოლო **Y-axis**(დერძზე) – **Observed**(დაკვირვება).

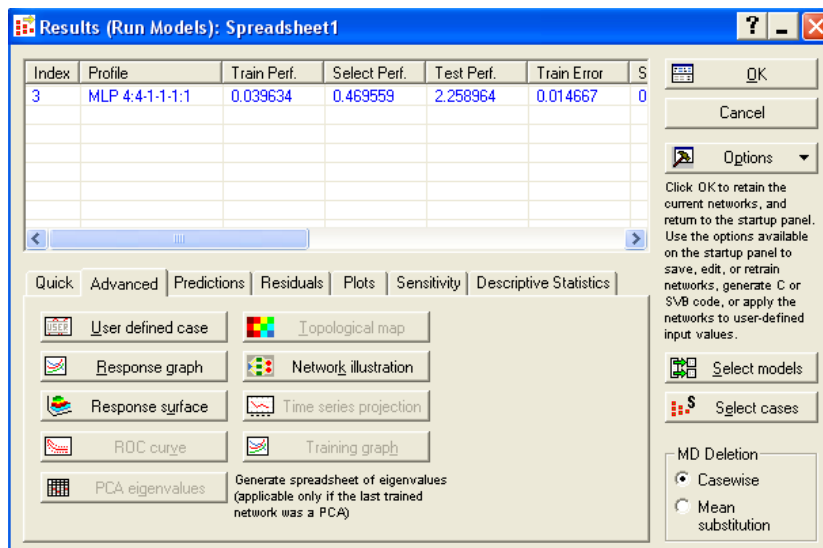


ავგოთ X და Y ცვლადების დამოკიდებულების გრაფიკი. ამისათვის **Click Graph X versus Y** ღილაკზე. ეკრანზე გამოდის დამოკიდებულების გრაფიკი

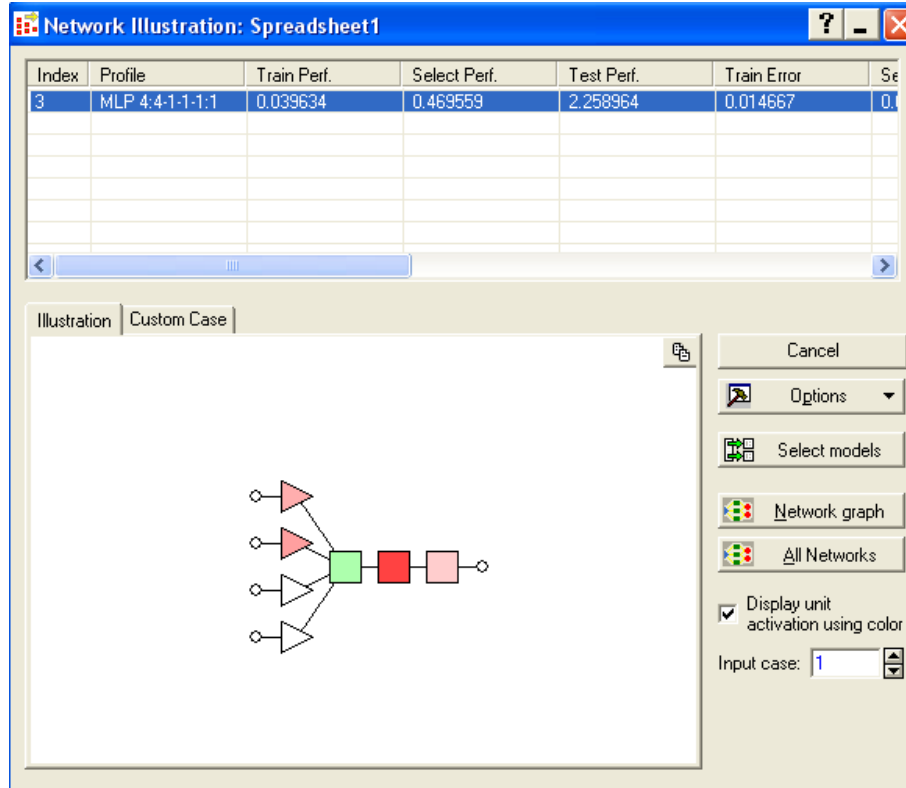


საიდანაც ჩანს, რომ გრაფიკი ფაქტიურად იდეალურად დევს სწორ ხაზზე. ამ შედეგის მიღება მოსალოდნელი იყო, რადგან როგორც შედეგების **Results(Run Models)** ფანჯარაში ჩანს მე-5 მოდელის სწავლების შეცდომა შეადგენს მხოლოდ 0,15%.

ნეირონული ქსელის არქიტექტურის ასაგებად შედეგების **Results(Run Models)** ფანჯარაში ჩავრთოდ **Advanced** ჩანართი



და **Click Network Archition**(ქსელის არქიტექტურა)-ის ღილაკზე. ეკრანზე გამოდის ნეიროქსელის გრაფი.



ქსელს აქვს პირდაპირი გავრცელების სტრუქტურა ოთხი შემავალი ცვლადით, ერთი ფარული შრით, რომელიც შეიცავს ერთ ელემენტს და ერთ ერთელემენტიან გამოშვებულ შრეს. მოდელის აგების სიზუსტე, ტესტირების ეფექტიანობის **TestPerf.** მაჩვენებლიდან გამომდინარე, 2,25% ტოლოა, რაც არასაკმარისი შედეგია.

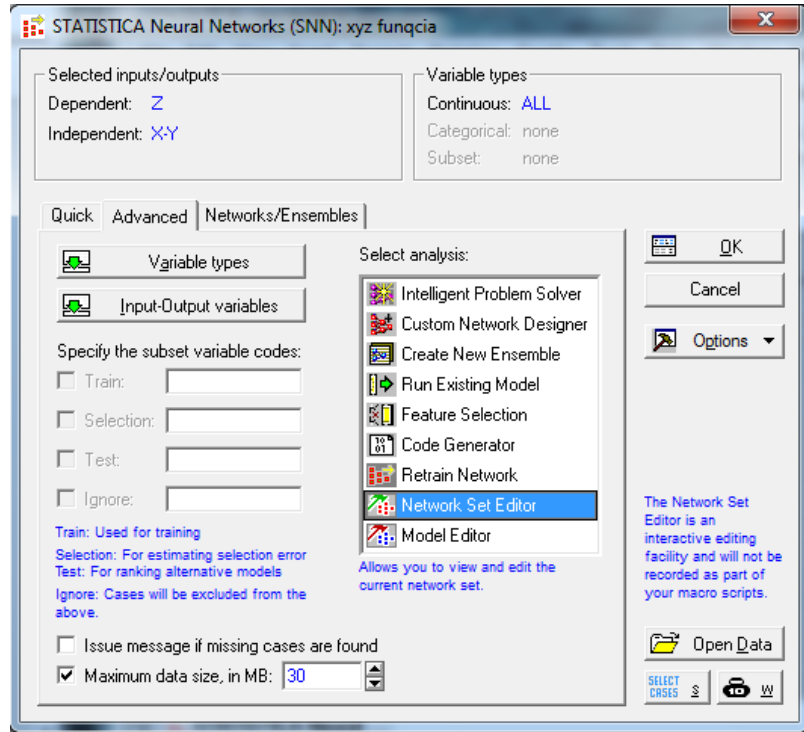
იმისათვის, რომ საბოლოოდ დაერწმუნდეთ ნეიროქსელის შერჩევის სისწორეში **Results(Run Models** ფანჯარაში ჩაერთოდ **Predictions** ჩანართი და **Click Predictions** ღილაკზე. ეკრანზე გამოდის შედეგების ცხრილი,

Prediction (3) (Spreadsheet1)		
	z	z.3
1	1357,921	1652,803
2	3601,639	3613,136
3	-0,491	-150,879
4	148,892	163,665
5	98,318	-54,211
6	87,441	602,561
7	0,061	-215,899
8	0,076	182,492
9	93,155	-65,686
10	-0,086	-201,365
11	-0,393	41,302
12	8472,709	8404,828
13	0,073	-111,661
14	0,234	-59,584

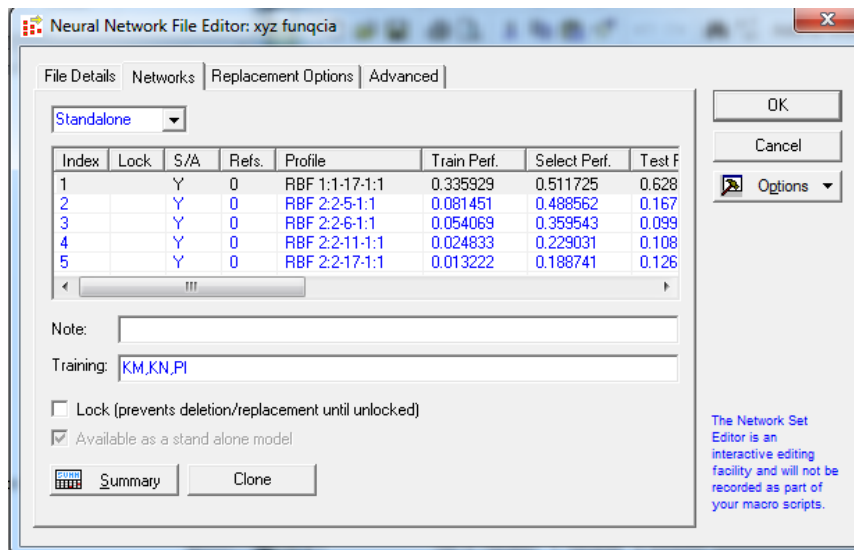
სადაც მარცხენა სვეტში მოცემულია z -ის ჭეშმარიტი მნიშვნელობები, ხოლო მარჯვენა სვეტში – ნეიროქსელის მიერ პროგნოზირებული მნიშვნელობები. როგორც ცხრილიდან ჩანს, მიღებული შედეგები არადაამაკმაყოფილებელია. განსაკუთრებით ცუდი შედეგია, როცა z ღებულობს ნულთან მიახლოებულ მნიშვნელობებს. ამ შემთხვევაში ცდომილება საკმაოდ დიდია. ასევე დიდია ცდომილება, როცა $z < 0,2$ სიდიდეზე. როცა $z > 100$, მაშინ შედეგი ასე თუ ისე დამაკმაყოფილებელია.

ბიჯი 8. ქსელის რედაქტირება.

STATISTICA Neural Networks სასტარტო ფანჯარაში ჩავრთოდ *Advanced* (დამატებითი) ჩანართი

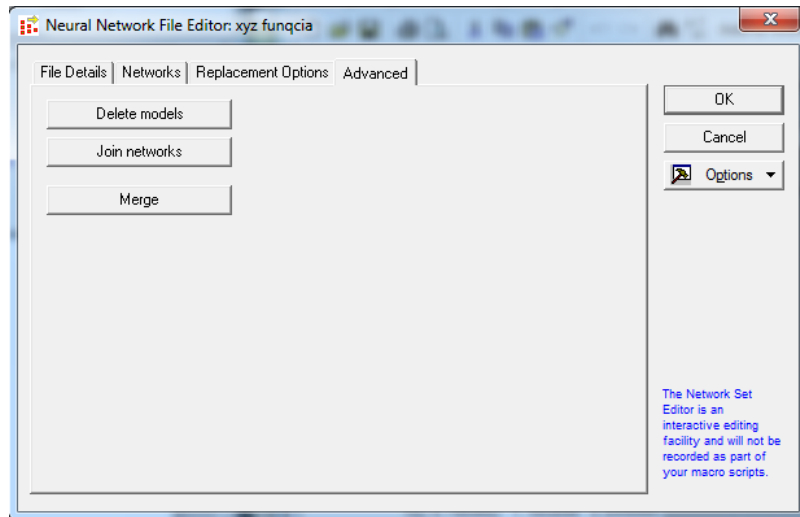


და **Select analysis** ველში მოვნიშნოთ **Network Set Editor**. შემდეგ **Click OK**. ეკრანზე გამოდის **Neural Networks File Editor** ფანჯარა.

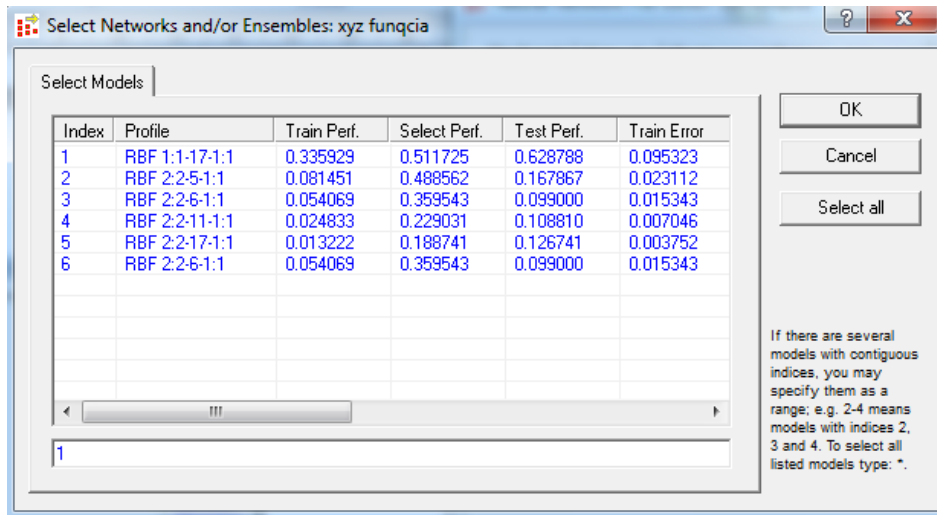


ჩავრთოდ **Networks** ჩანართი. ამ დიალოგიურ ფანჯარაში შეიძლება დავათვალიეროთ, ამოვირჩიოთ და დავაღაგოთ სხვადასხვა ნეირონული ქსელები და ანსამბლები.

ქსელის შესწავლის შემდეგ შეგვიძლია წავშალოთ ზოგიერთი მათგანი. ქსელის წასაშლელად **Neural Networks File Editor** ფანჯარაში ჩავრთოდ **Advanced** ჩანართი და



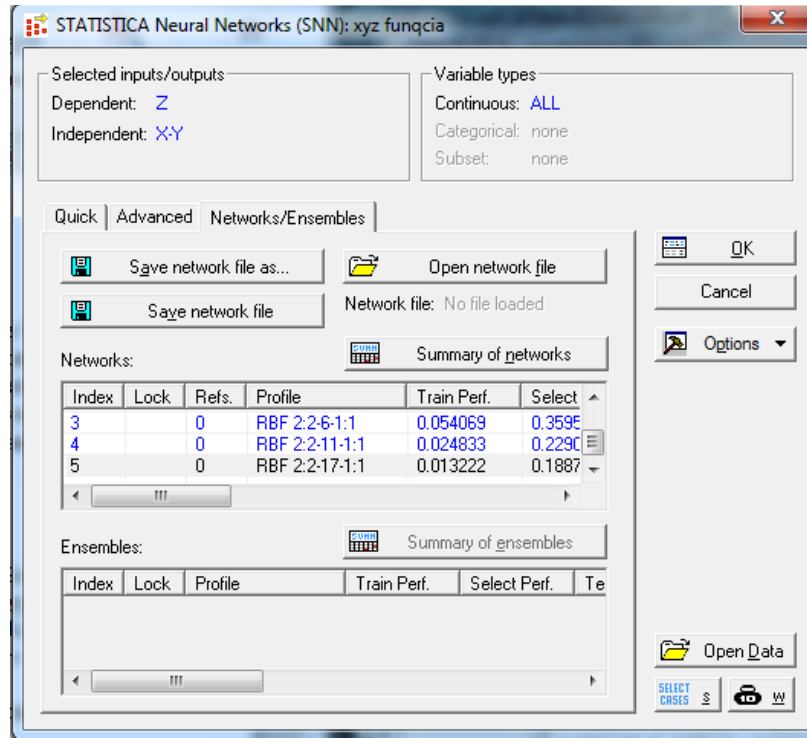
ეკრანზე გამოსულ ფანჯარაში *Click Delete models* ღილაკზე. ეკრანზე გამოდის *Select Networks and/or Ensembles* ფანჯარა.



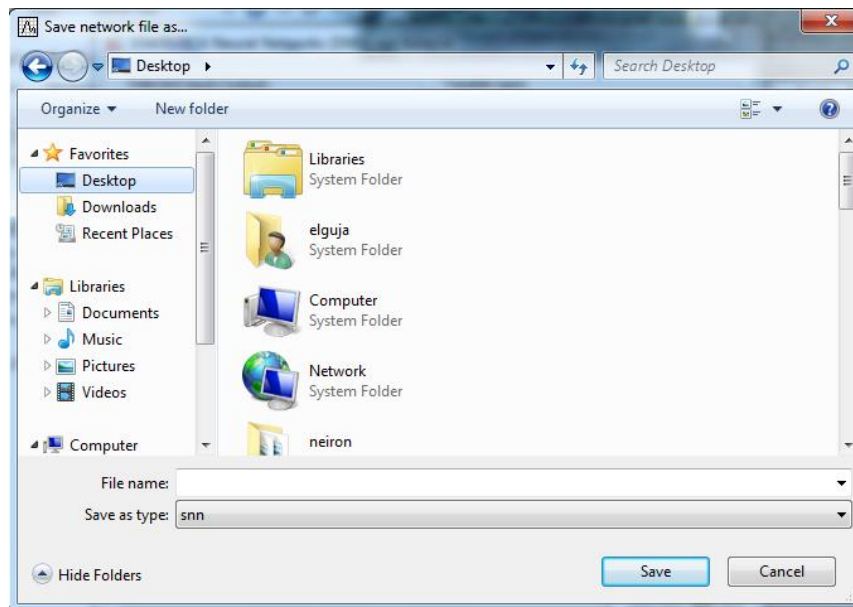
მოვნიშნოთ წასაშლელი ქსელი ან ქსელები და შემდეგ *Click OK*.

ბიჯი 9. ნეიროქსელის შენახვა.

. ნეირონული ქსელების შესანახად ვბრუნდებით სასტარტო ფანჯარაში. ჩაერთოდ *Networks/Ensembles* (ქსელები / ანსამბლები) ჩანართი



Click **Save network file as..** ღილაკზე. ეკრანზე გამოდის ამავე დასახელების **Save network file as..** ფანჯარა



სადაც **File name** ველში ჩაეწეროთ შესაბამისი ფაილის სახელი (ფაილს აქვს გაფართოება *.snn*). შევარჩიოთ ფაილის შენახვის ადგილი (მაგალითად, *Desktop*, *Local Disk (C)* და სხვა). შემდეგ **Click Save** ღილაკზე.

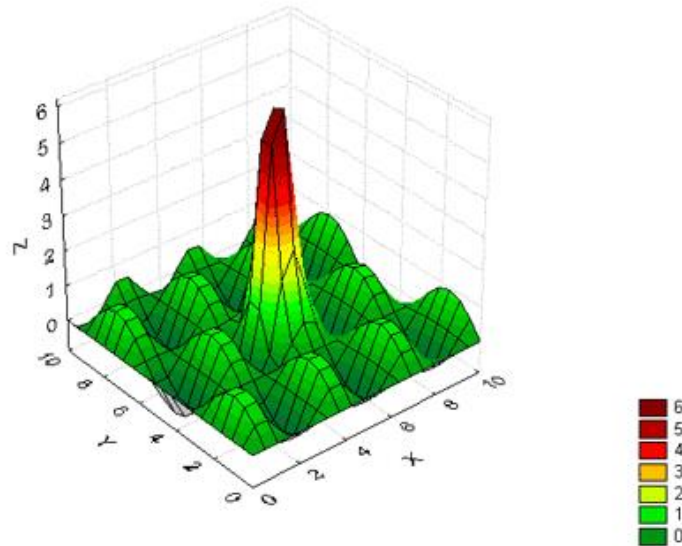
დაახლოებით ასეთი პროცედურით ხდება **STATISTICA 7** სისტემაში ნეირონული ქსელებით სხვადასხვა ტიპის ამოცანების გადაწყვეტა.

პრაქტიკული სამუშაო 2

ფუნქციის აპროქსიმაცია

მოცემულია ორი ცვლადის უცნობი ფუნქცია, რომელიც წარმოდგენილი არიან 1000 მნიშვნელობებით ინტერვალში $[0,10] \times [0,10]$. იგულისხმება, რომ წერტილები, სადაც მოცემულია ფუნქციის მნიშვნელობები, მოცემულ კვადრატში განაწილებულია შემთხვევით.

უნდა ავაგოთ მათემატიკური მოდელი, რომელიც აღწერს დამოკიდებულებას ფუნქციის მნიშვნელობასა და დამოუკიდებელი ცვლადების მნიშვნელობებს შორის. ფუნქციის გრაფიკული გამოსახულება მოცემულია შემდეგ ნახაზზე:

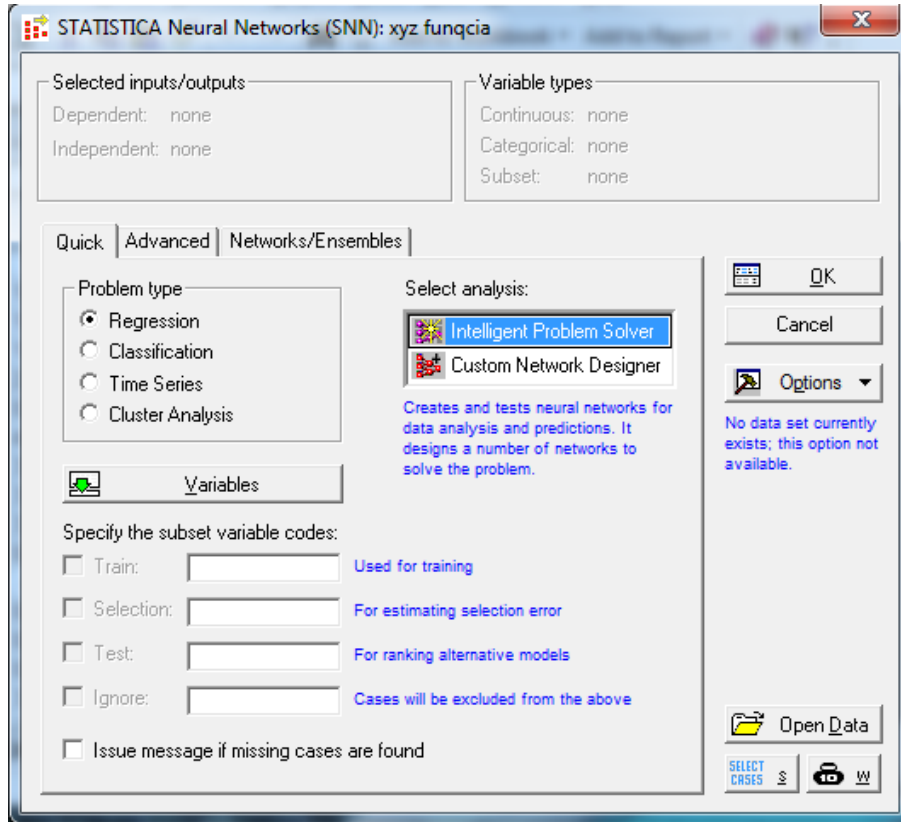


გავხსნათ *xyzfunctia*. მონაცემთა ფაილი, რომელიც შედგება სამი ცვლადისაგან: X, Y, Z . X და Y –დამოუკიდებელი ცვლადებია, ხოლო ცვლადი Z –დამოკიდებული.

	1 X	2 Y	3 Z
1	1,56684	2,071825	2,36023
2	1,221973	0,930952	0,250917
3	1,031099	2,775307	4,035066
4	1,848694	3,446776	4,969204
5	2,427282	1,908399	1,603156
6	0,744185	1,327998	1,283903
7	4,036675	3,460967	3,903596
8	3,902029	2,973617	2,996219
9	3,549554	2,938368	3,101959
10	3,763685	4,323054	5,764265
11	0,282497	0,471174	-0,19777
12	4,19236	3,197317	3,298455
13	0,891796	2,766762	4,087627
14	0,624997	1,536103	1,759708
15	0,785442	1,145894	0,899066
16	3,794871	4,010025	5,122615
17	0,372507	2,497742	3,80923
18	1,85062	2,814768	3,704226
19	2,363822	3,158815	4,13572
20	4,556097	3,430906	3,583764
21	0,80317	1,235618	1,069651
22	8,818658	6,858928	8,308527
23	1,870483	2,274316	2,61339
24	1,740074	2,479156	3,088275
25	4,156881	3,09389	3,10934
26	0,123189	0,316279	-0,42904
27	8,858022	7,122111	8,815211
28	8,113952	4,49707	3,937164
29	3,111362	1,004074	-0,54753
30	8,767844	4,32048	2,74288

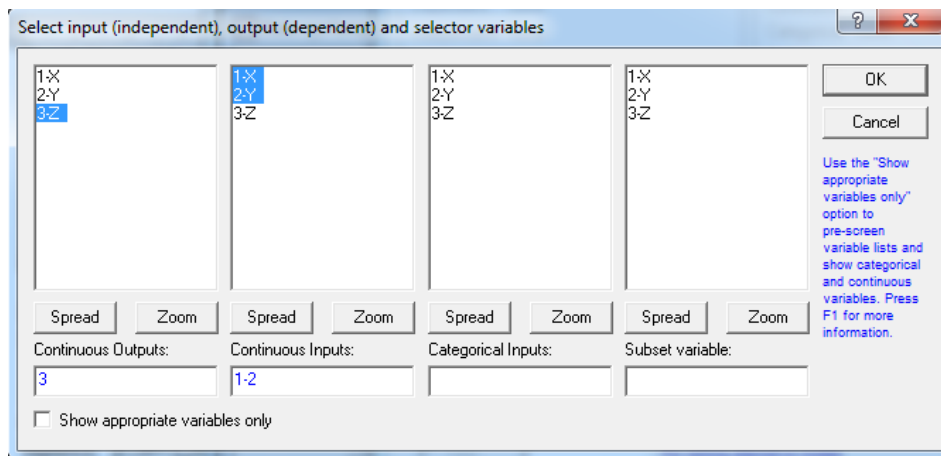
.ამოცანის ამოხსნის კლასიკური მიდგომა მდგომარეობს იმაში, რომ პირველ ბიჯზე ვაგებთ მრავლობითი რეგრესიის მოდელს. აშკარაა, რომ დამოკიდებულება არაწრფივია და ამიტომ ყველა წრფივი მოდელიები მოგვცემენ "ცუდ" შედეგს. ამ მაგალითისთვის ავაგოთ ნეიროქსელური მოდელი.

სასტარტო პანელის **STATISTICA Neural Networks** გამოსაძახებლად, მენიუდან **Statistics** შევარჩიოთ ბრძანება **Neural Networks**. ეკრანზე გამოდის **STATISTICA Neural Networks(SNN)** ფანჯარა,

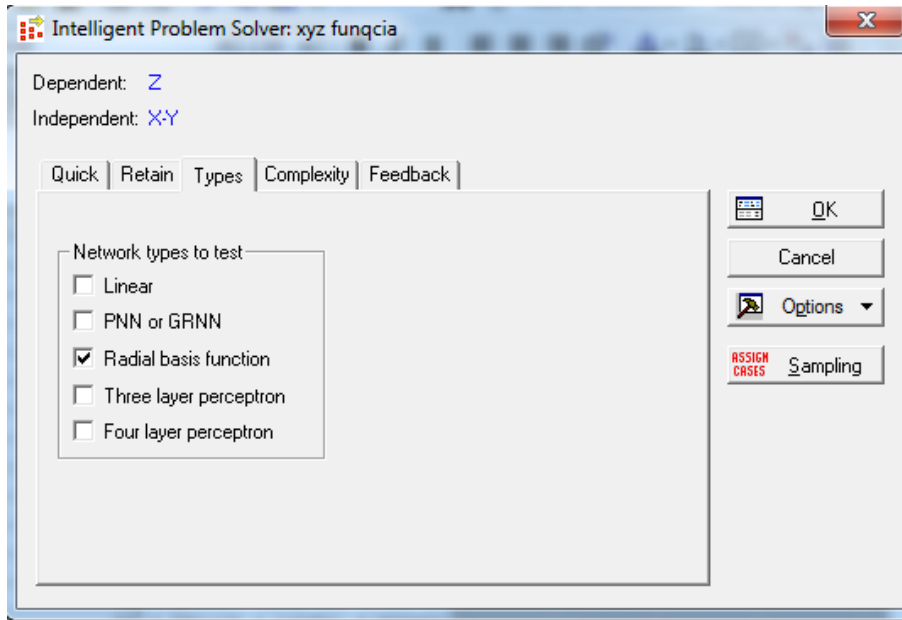


სადაც ჩაერთოდ **Quick** ჩანართი და **Problem type** ველში შევარჩიოთ ამოცანის ტიპი **Regression**, ხოლო **Select analysis** ველში ამოხსნის მეთოდი – **Intelligent Problem Solver** (გადაწყვეტის ოსტატი).

ცვლადების მოსანიშნად **Click Variables** ღილაკზე და ეკრანზე გამოსულ ფანჯარაში მოვნიშნოთ **Continuous Outputs**:(უწყვეტი გამომავალი) ველში Z ცვლადი, ხოლო **Continuous Inputs**(უწყვეტი შემავალი) ველში – x,y ცვლადები. შემდეგ **OK**.

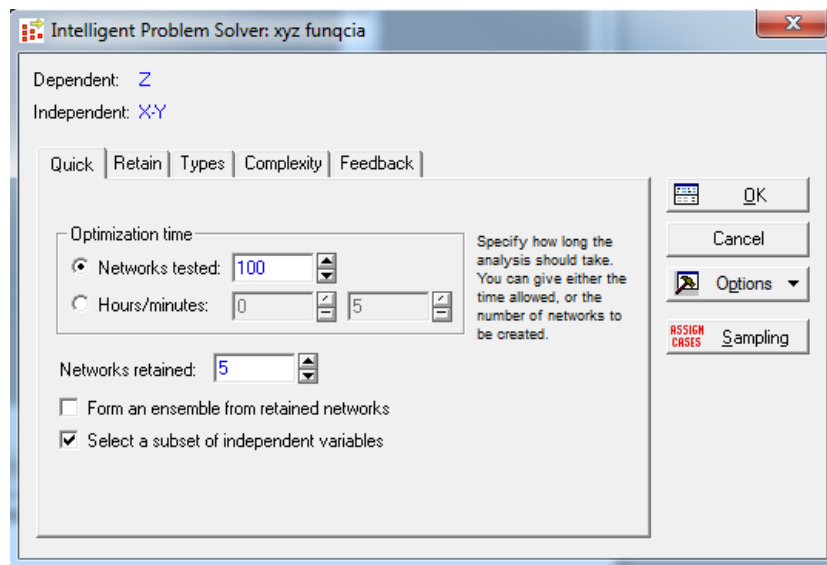


ვბრუნდებით *STATISTICA Neural Networks (SNN)* ფანჯარაში და კვლავ *OK*.
 ეკრანზე გამოდის *Intelligent Problem Solver*(გადაწყვეტის ოსტატი)-ის ფანჯარა.
 ქსელის ტიპის შესარჩევად ჩავრთოდ *Types*(ტიპი) ჩანართი.

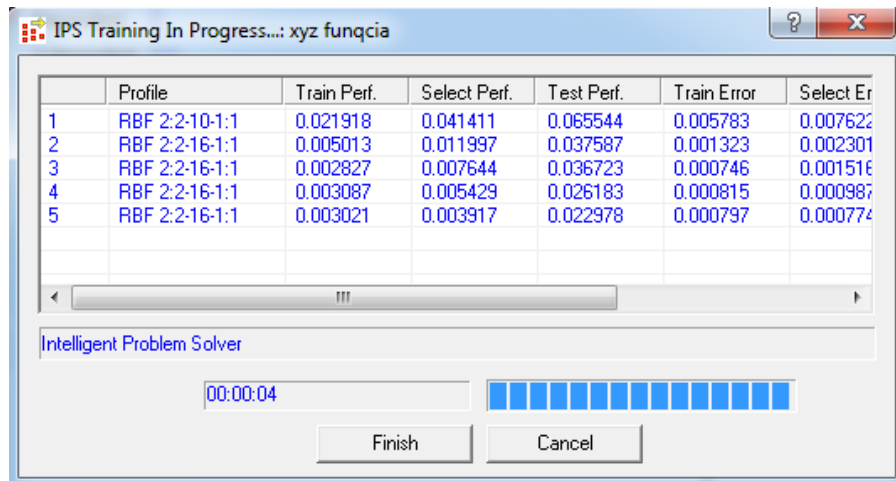


როგორც ჩანს ეს ამოცანა ტოპოლოგიურად ახლოს არის ისეთ ნეიროქსელურ არქიტექტურასთან, რომელიც ეფუძნება რადიალურ-ბაზისურ ფუნქციებს, ამიტომ *Network types to test* ველში ჩავრთოდ *Radial basis function* (რადიალური ბაზისური ფუნქცია)-ის ოპცია.

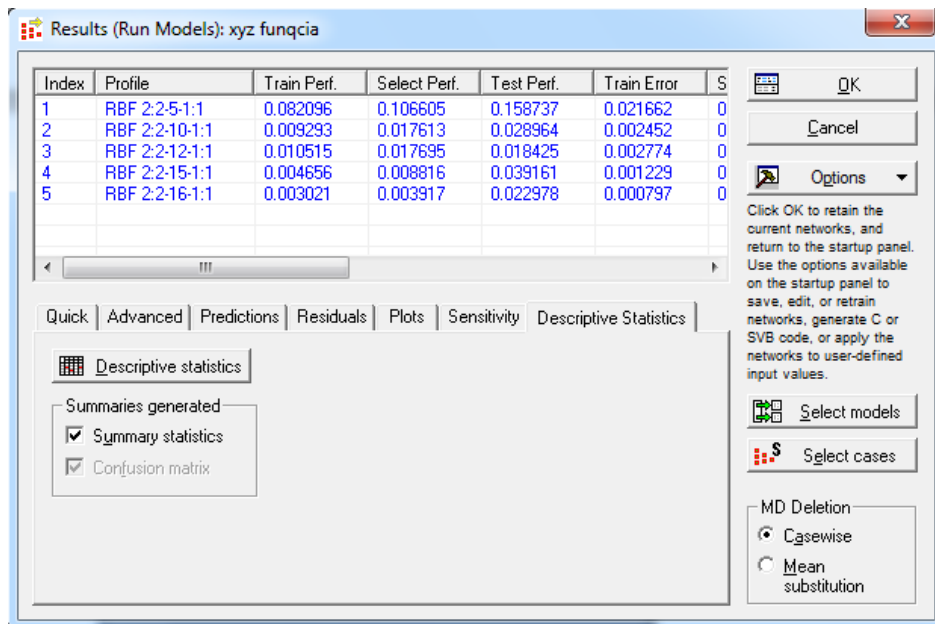
ჩავრთოდ *Quick*(სწრაფი) ჩანართი,



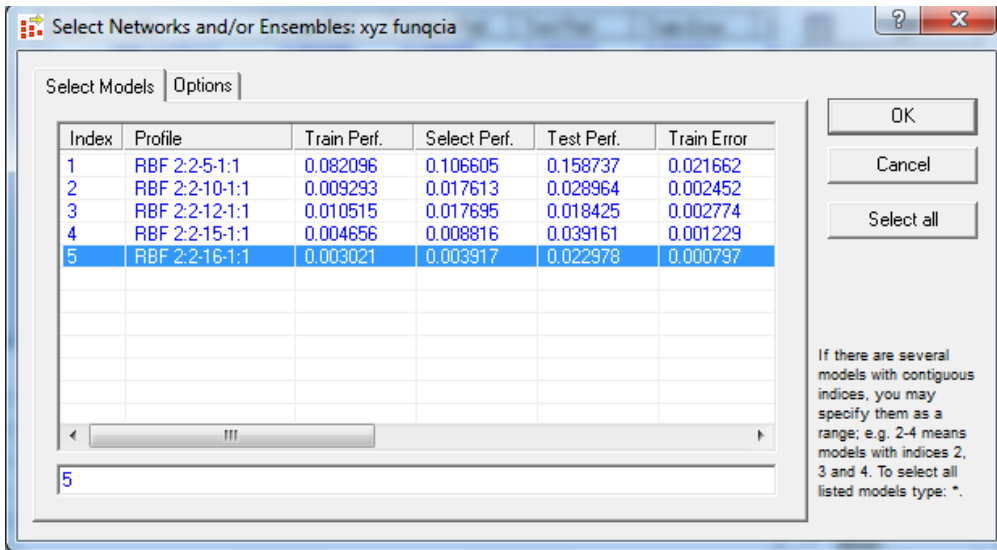
სადაც **Optimization time** (ანალიზის ხანგრძლივობა) ჯგუფის **Networks tested** (ქსელების ტესტირება) ველში ჩავეწერთ 100, ხოლო ქსელის შესანახად **Networks retained** (ქსელის შენარჩუნება) ველში ჩავეწერთ ციფრი 5, რათა პროგრამამ შეინახოს 5 საუკეთესო ქსელის ვარიანტი. შემდეგ **Click OK** ღილაკზე. ეკრანზე გამოჩნდება ნეიროქსელის სწავლების დიალოგიური **Training In Progress**(მიმდინარეობს შესწავლა) ფანჯარა.



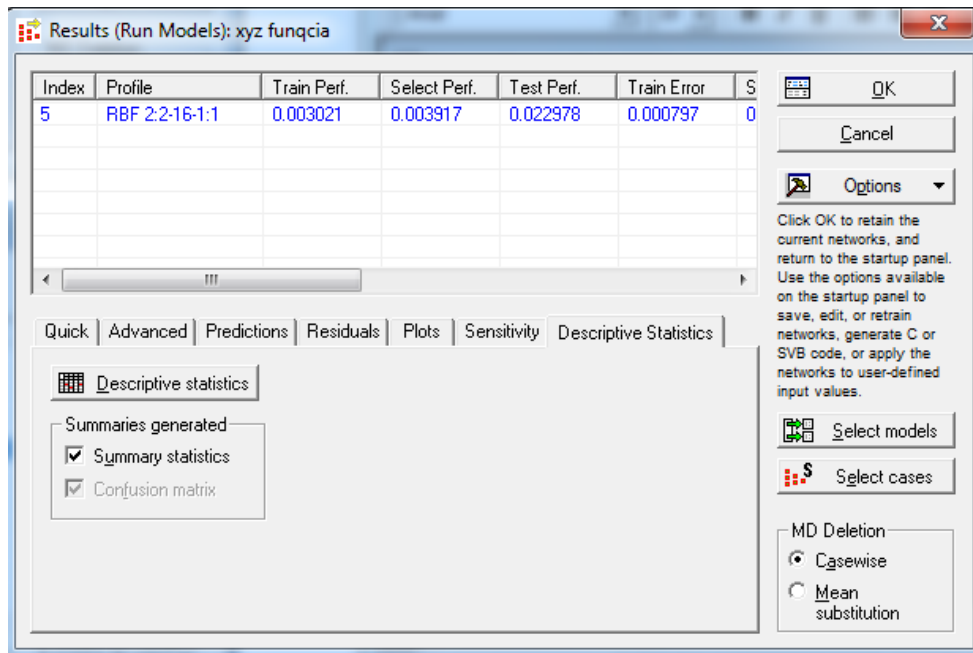
ძეზის დასრულების შემდეგ ეკრანზე გამოჩნდება შედეგების **Results(Run Models)** (შედეგები) ფანჯარა.



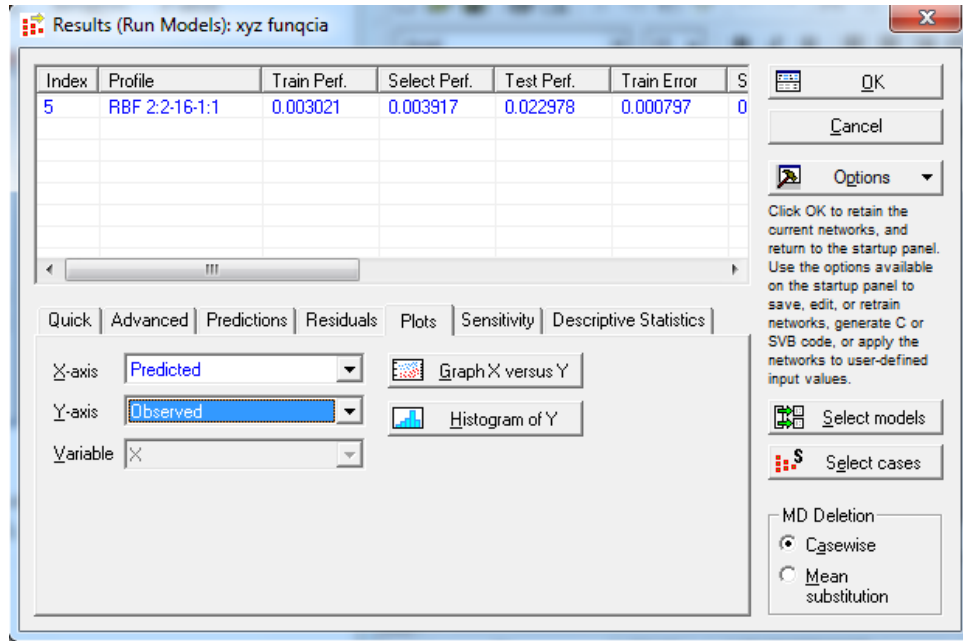
Click Select models (მოდელის მონიშვნა) ღილაკზე და ეკრანზე გამოსულ შედეგების ცხრილში მონიშნეთ მე-5 მოდელი, რადგან მას გააჩნია ყველაზე კარგი შედეგი.



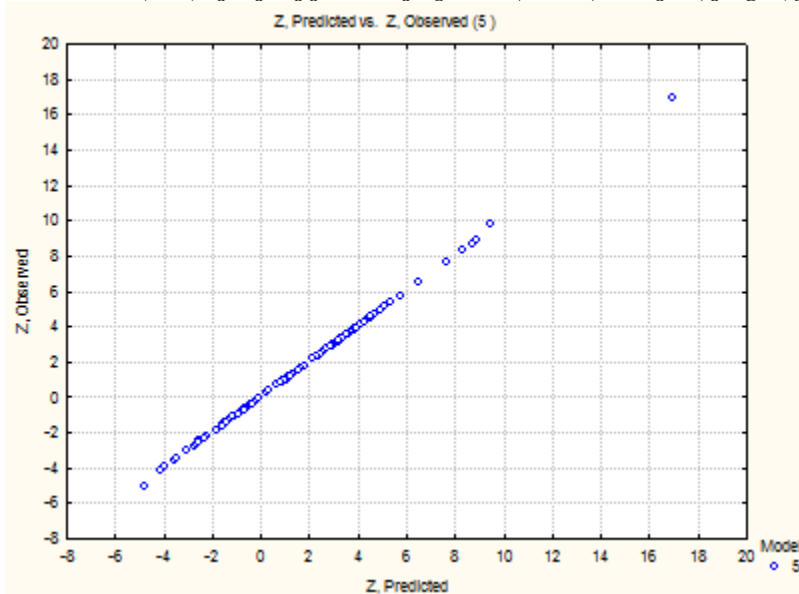
შეძლებს *OK*.



დაკვირვებების გრაფიკის ასაგებად ჩაერთოდ ჩანართ *Plots* (გრაფიკი) და *X-axis* (დერძზე)-ის ველში შევარჩიოდ *Predicted* (დაკვირვება), ხოლო *Y-axis* (დერძზე) – *Observed* (პროგნოზირება).



აგვით X და Y ცვლადების დამოკიდებულების გრაფიკი. ამისათვის Click **Graph X versus Y** ლილაკზე. ეკრანზე გამოდის დამოკიდებულების გრაფიკი,



საიდანაც ჩანს, რომ გრაფიკი ფაქტიურად იდეალურად დევს სწორ ხაზზე. ამ შედეგის მიღება მოსალოდნელი იყო, რადგან როგორც შედეგების **Results(Run Models)** ფანჯარაში ჩანს მე-5 მოდელის სწავლების ცდომილება შეადგენს მხოლოდ 0,08%.

შედეგების **Results(Run Models)** ფანჯარის **Quick** ჩანართში Click **Descriptive statistics** (აღწერითი სტატისტიკა)-ის ლილაკზე. ეკრანზე გამოდის რეგრესიის შეცდომების ცხრილი,

	Regression (5) (xyz funqcia)		
	Z.5		
Data Mean	1,654097		
Data S.D.	3,417326		
Error Mean	-0,001730		
Error S.D.	0,040803		
Abs E. Mean	0,014707		
S.D. Ratio	0,011940		
Correlation	0,999932		

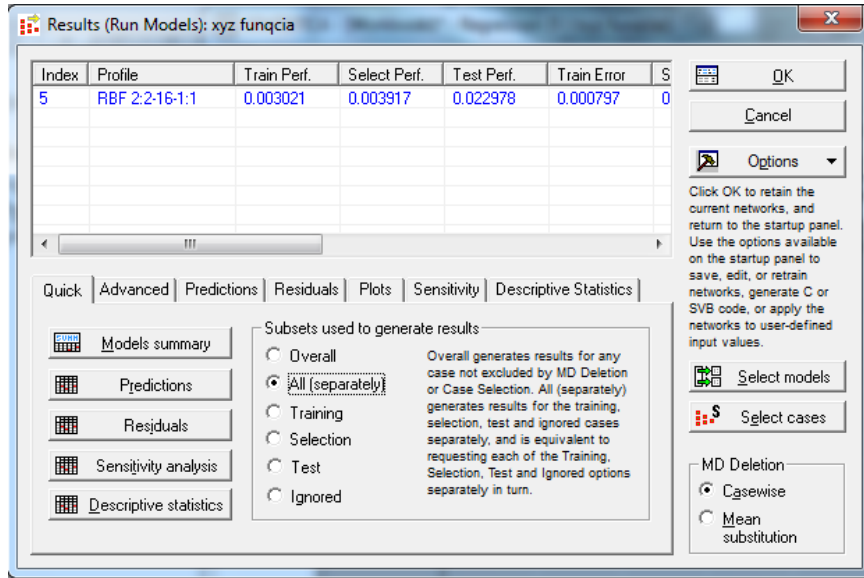
მიღებულ შედეგებიდან ჩანს, რომ მე-5 ქსელის სტანდარტული გადახრების ფარდობის სიდიდე **S.D.Ratio** 0,0119 ტოლია, რაც საკმაოდ კარგი მაჩვენებელია.

თუ შედეგების **Results(Run Models)** ფანჯარის **Subets used to generale resultsc** (შერჩევის შედეგების გამოყვანა)-ის ველში **Overall** (სრული) ოპციის მაგივრად ჩავრთავთ **All (separately)** (ყველა ცალკე) ოპციას, მაშინ ეს მოგვცემს საშუალებას მივიღოთ ცხრილი, სადაც შერჩევის შედეგები ყველა ამონარჩევისათვის ცალკე გამოდის.

	Regression (5) (xyz funqcia)			
	T.Z.5	S.Z.5	X.Z.5	I.Z.5
Data Mean	1,858886	1,595856	1,287589	0,00
Data S.D.	3,722170	2,546201	3,483434	0,00
Error Mean	0,000000	-0,004443	-0,002606	0,00
Error S.D.	0,011245	0,009974	0,080042	0,00
Abs E. Mean	0,008911	0,007468	0,033965	0,00
S.D. Ratio	0,003021	0,003917	0,022978	0,00
Correlation	0,999995	0,999993	0,999783	0,00

როგორც ცხრილიდან ჩანს **S.D.Ratio** სიდიდე ყველა ამონარჩევისათვის თითქმის ერთნაირია.

იმისათვის, რომ საბოლოოდ დავრწმუნდეთ ნეიროქსელის შერჩევის სისწორეში **Results(Run Models)** ფანჯარაში ჩავრთოდ **Predictions** ჩანართი.

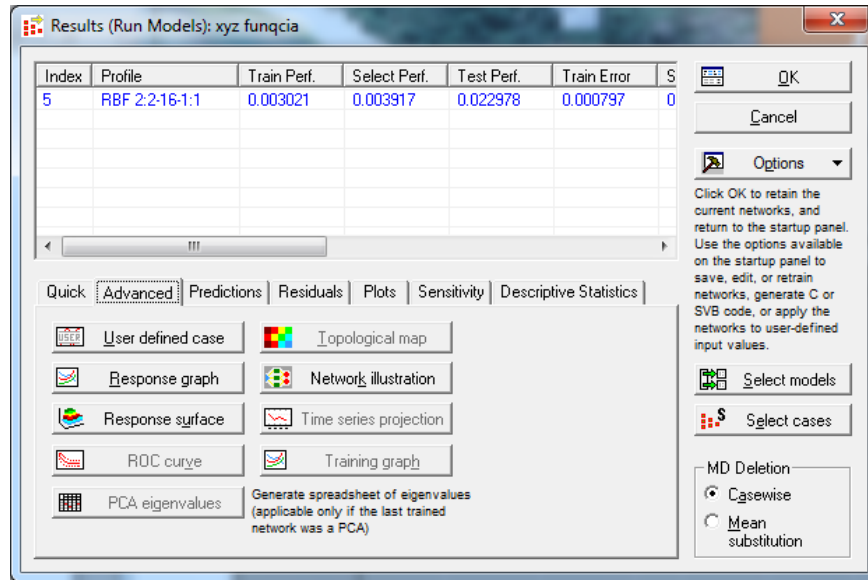


და Click **Predictions** (პროგნოზირება)-ის ღილაკზე. ეკრანზე გამოდის შედეგების ცხრილი,

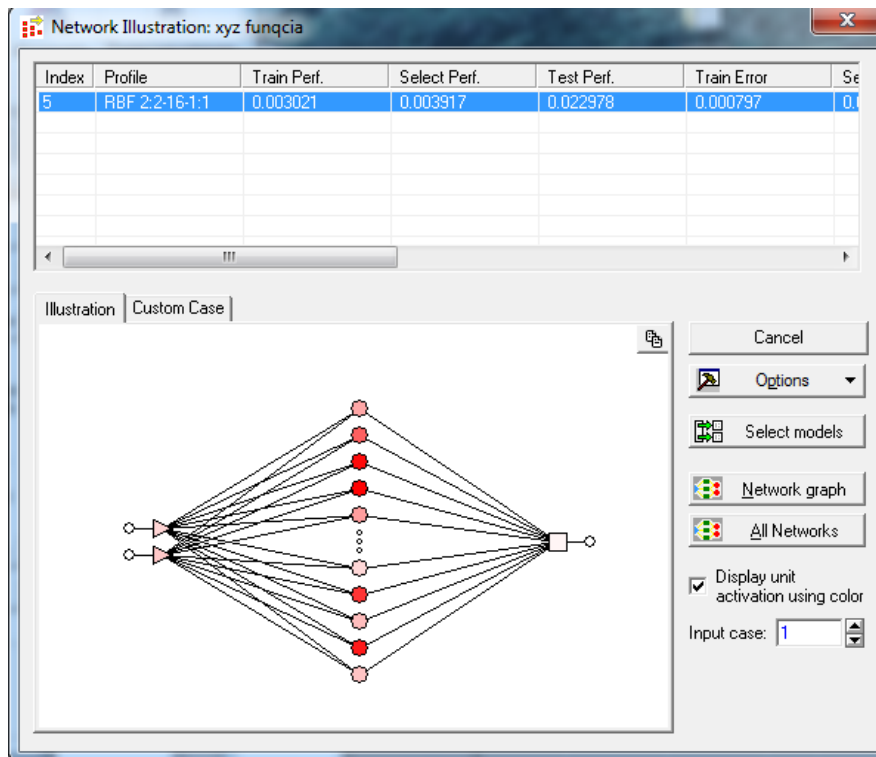
	Prediction (5) (Test) (xyz funqcia)	
	Z	Z.5
1	2,36023	2,35839
7	3,90360	3,89803
12	3,29845	3,30537
24	3,08827	3,08128
25	3,10934	3,11988
29	-0,54753	-0,54004
37	-2,49322	-2,52592
38	-2,19896	-2,22933
40	1,01722	1,02118
44	-0,36025	-0,34414
53	8,67305	8,74829
63	9,79347	9,45707
64	-0,06371	-0,06382
71	1,17087	1,19267
72	-3,46453	-3,47633
74	2,96469	2,95817
75	-3,44946	-3,47298
76	5,32762	5,35440
77	-1,51580	-1,53374
78	-0,26789	-0,25244

სადაც მარცხენა სვეტში მოცემულია z-ის ჭეშმარიტი მნიშვნელობები, ხოლო მარჯვენა სვეტში – ნეიროქსელის მიერ ტესტური ამონარჩევით მიღებული პროგნოზირებული მნიშვნელობები. როგორც ცხრილიდან ჩანს, მიღებული შედეგები დამაკმაყოფილებელია.

ნეირონული ქსელის არქიტექტურის ასაგებად შედეგების **Results(Run Models)** ფანჯარაში ჩაერთოდ **Advanced** ჩანართი



და **Click Network Archition** (ქსელის არქიტექტურა)-ის ღილაკზე. ეკრანზე გამოდის ნეიროქსელის გრაფი.



ქსელს აქვს პირდაპირი გავრცელების სტრუქტურა ორი შემავალი ცვლადით, ფარული შრე 16 ელემენტით და ერთი ელემენტი გამოშვალ შრეში. მოდელის აგების სიზუსტე, ტესტირების ეფექტიანობის **TestPerf.** მაჩვენებლიდან გამომდინარე, 0,08% ტოლოა.

შეიძლება გამოვიტანოთ დასკვნა, რომ ქსელმა წარმატებით მოახდინა შესასწავლი ნაკრების სტრუქტურის გარჩევა და მისი გამოყენება შეიძლება დამოკიდებული ცვლადის მნიშვნელობების პროგნოზირებისთვის. შევინახოთ მოდელი. ისე როგორც მოცემულია პრაქტიკული სამუშაო 1 –ში

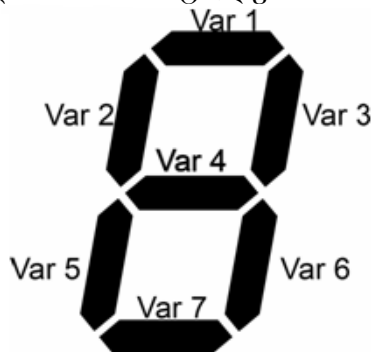
პრაქტიკული სამუშაო 3

სახეთა ბარჩევის ამოცანა

ნეირონულ ქსელების საშუალებით შესაძლებელია სახეთა გარჩევის ამოცანის გადაწყვეტა. მაგალითისათვის განვიხილოთ იმიტირებული მონაცემები, წარმოადგენილი ციფრების სახით, რომლებიც გამონათებულნი არიან გაუმართავი კალკულატორის ეკრანზე. დამოკიდებულ **Digit** ცვლადის დაკვირვების კლასები შეესაბამებიან ციფრებს 0, 1, 2, 3, .. 9, რომლებიც შეყვანილია კალკულატორის კლავიშიდან.

ამოცანაში გვაქვს 7 კატეგორიული პრედიქტორები **Var1 - Var7** ხაზების სახით, რომლებიც ქმნიან ციფრების კონფიგურაციას. კატეგორიული პრედიქტორის დონე (0 – არ არის; 1 - არის) გვიჩვენებს, გაანათა თუ არა ეკრანზე შესაბამისმა ხაზმა იმ შეიდი ხაზიდან (3 ჰორიზონტალური და 4 ვერტიკალური), რომლითაც წარმოდგენილია გასარჩევი ციფრი.

პრედიქტორის ცვლადების აღწერა : **Var1**-ზედა ჰორიზონტალური, **Var2**–ზედა მარცხენა ვერტიკალური, **Var3** -ზედა მარჯვენა ვერტიკალური, **Var4** –შუა ჰორიზონტალური, **Var5** –ქვედა მარცხენა ვერტიკალური, **Var6** –ქვედა მარჯვენა ვერტიკალური და **Var7**–ქვედა ჰორიზონტალური.



კალკულატორი გაუმართავია, ასე რომ, როდესაც დავაჭერთ ციფრულ კლავიატურის ნებისმიერ ღილაკს, ეკრანზე ყოველთვის არ გამოანათებს შესაბამისი ხაზების კომბინაცია.

Statistica სისტემის **Datasets** ბიბლიოთეკიდან გავხსნათ **Digit** მონაცემთა ფაილი, რომლის ფრაგმენტი წარმოდგენილია შემდეგ ნახაზზე:

Data: Digit (8v by 500c)								
Example data for pattern recognition								
	1	2	3	4	5	6	7	8
	DIGIT	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7
1	seven	ONE	ZERO	ONE	ZERO	ZERO	ONE	ZERO
2	one	ZERO	ZERO	ONE	ZERO	ZERO	ONE	ZERO
3	four	ZERO	ONE	ONE	ONE	ZERO	ONE	ZERO
4	two	ONE	ONE	ONE	ONE	ONE	ZERO	ZERO
5	eight	ZERO	ONE	ONE	ONE	ONE	ONE	ONE
6	one	ZERO	ZERO	ONE	ZERO	ZERO	ONE	ZERO
7	five	ONE	ONE	ZERO	ONE	ZERO	ONE	ONE
8	six	ONE	ZERO	ZERO	ONE	ONE	ONE	ONE
9	two	ONE	ZERO	ONE	ONE	ONE	ZERO	ONE
10	eight	ONE	ONE	ONE	ONE	ZERO	ONE	ONE
11	one	ZERO	ZERO	ONE	ZERO	ONE	ONE	ZERO
12	eight	ONE	ONE	ONE	ONE	ONE	ONE	ONE
13	one	ONE	ZERO	ONE	ZERO	ONE	ONE	ZERO
14	seven	ONE	ZERO	ONE	ZERO	ZERO	ONE	ZERO
15	seven	ZERO	ZERO	ONE	ZERO	ONE	ONE	ZERO
16	six	ONE	ONE	ZERO	ONE	ONE	ONE	ONE
17	zero	ONE	ZERO	ONE	ZERO	ONE	ONE	ZERO
18	one	ZERO	ZERO	ONE	ZERO	ZERO	ONE	ONE
19	four	ZERO	ONE	ONE	ONE	ZERO	ONE	ZERO
20	six	ONE	ONE	ZERO	ONE	ONE	ONE	ONE

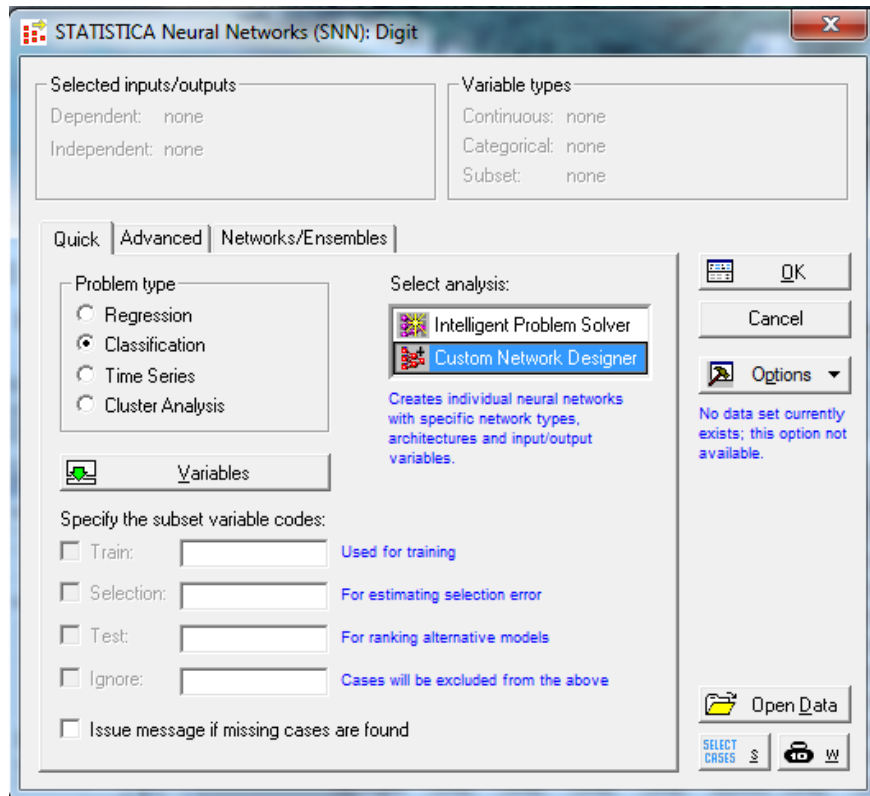
ფაილი წარმოდგენს მონაცემების მთელ კომპლექტს, რომელიც შედგება 500 დაკვირვებებისაგან. გარჩევის შედეგები ჩაწერილია **Digit** ცვლადის სვეტში. **Var1 - Var7** ცვლადების სვეტებში წარმოდგენილია დამოუკიდებელი კატეგორიული პრედიქტორების დონეები (ნული ან ერთი).

მოცემულ მაგალითში საჭიროა ავავოთ რეალური კალკულატორისაგან მიღებული ციფრების გარჩევის მოდელი. მიზანშეწონილია ნეირონული ქსელის მიმართ ჩამოვყალიბოთ შემდეგი მოთხოვნები: 1) ქსელმა უნდა შეძლოს მონაცემთა ექსტრაპოლირება ანუ ჩამოაყალიბოს სწორი პროგნოზი პრედიქტორების კომბინაციებზე, რომლებიც ძალიან განსხვავდებიან სასწავლო სიმრავლისაგან. 2) პრაქტიკული გამოყენებისათვის პროგნოზირების მცირე დრო.

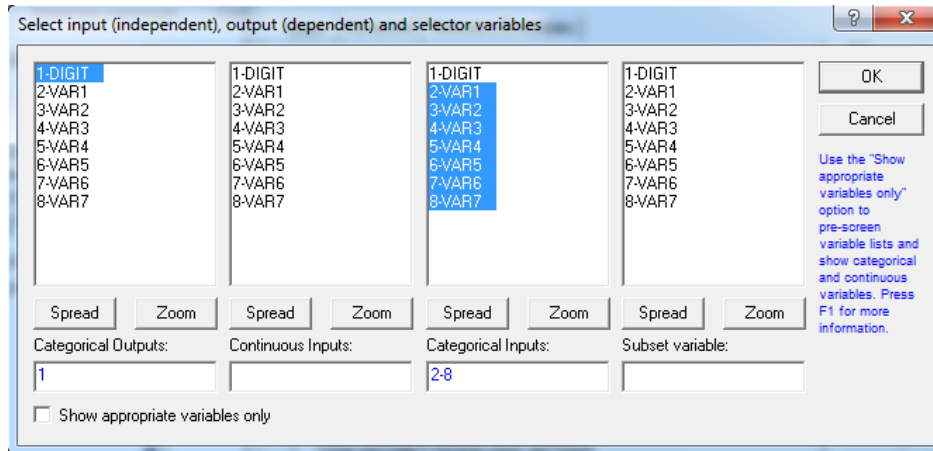
მითითებულ პირობებს შეესაბამება მრავალშრიანი პერსეპტრონის არქიტექტურა. ელემენტების რაოდენობა ფარულ შრეში შევარჩიოთ 5(კლასიფიკაცია წარმოებს 10 კლასსში). საზოგადოდ, მოდელის აგებისას პირველი მიახლოებით ფარული შრის ელემენტებს ირჩევენ შემავალი და

გამომავალი ელემენტების ჯამის ნახევრის სიდიდის ტოლად, მაგრამ ჩვენ შემთხვევაში ელემენტების რაოდენობის შერჩევა განპიროვნებულია კლასების რაოდენობითაც.

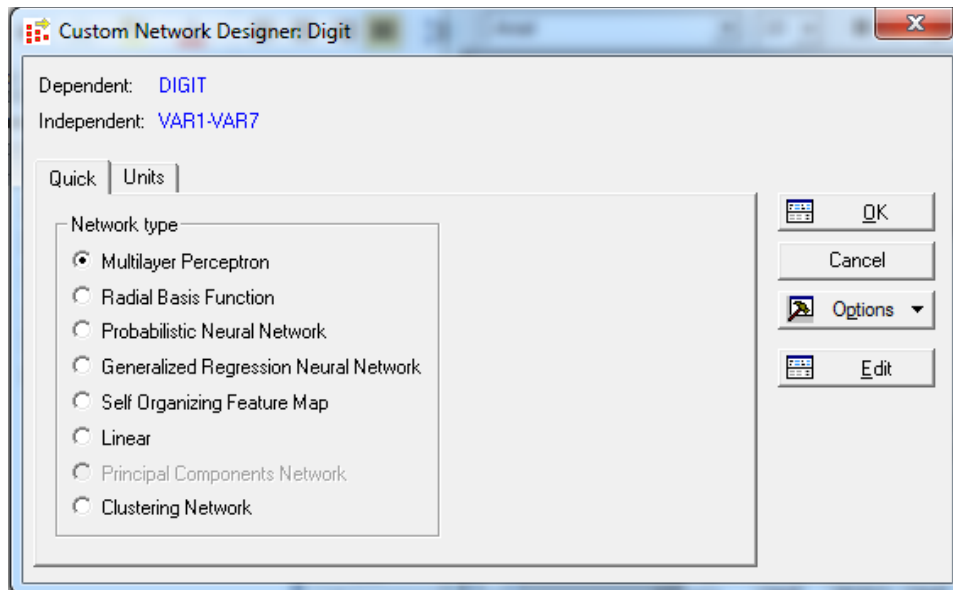
სასტარტო პანელის **STATISTICA Neural Networks** გამოსაძახებლად, მენიუდან **Statistics** შევარჩიოთ ბრძანება **Neural Networks**. ეკრანზე გამოდის **STATISTICA Neural Networks(SNN)** ფანჯარა.



ჩაერთოდ **Quick** ჩანართი და **Problem type** ველში შევარჩიოთ ამოცანის ტიპი **Classification**, ხოლო **Select analysis** ველში ამოხსნის მეთოდი – **Custom Network Designer** (ქსელის კონსტრუქტორი). ცვლადების მოსანიშნად **Click Variables** ლიკაკზე და ეკრანზე გამოსულ **Select input(independent), outp(dependent) and Selector vaiiables** (მოფნიშნოთ შემაგავალი (დამოკიდებული), გამომავალი (დამოკიდებული) და დამაჯგუფებელი ცვლადები) ფანჯრის **Continuous outputs** (უწყვეტი გამომავალი) ველში მოფნიშნოთ **DIGIT** ცვლადი, ხოლო **Categorical inputs** (შემაგავალი გატეგორიალური) ველში – **VAR 1- VAR7** ცვლადები.

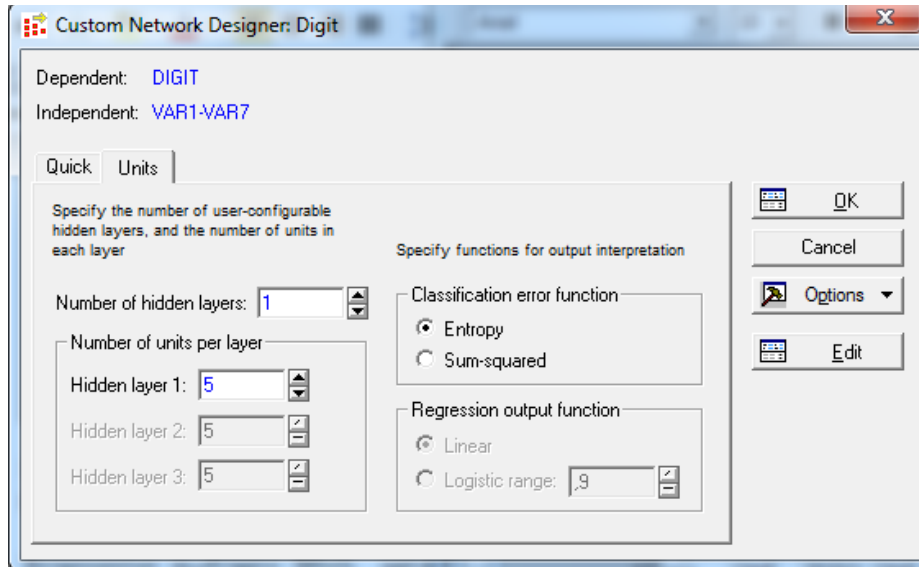


შემდეგ *Click OK*. ვბრუნდებით *STATISTICA Neural Networks (SNN)* ფანჯარაში და კვლავ *OK*. ეკრანზე გამოდის *Custom Network Designer* ფანჯარა

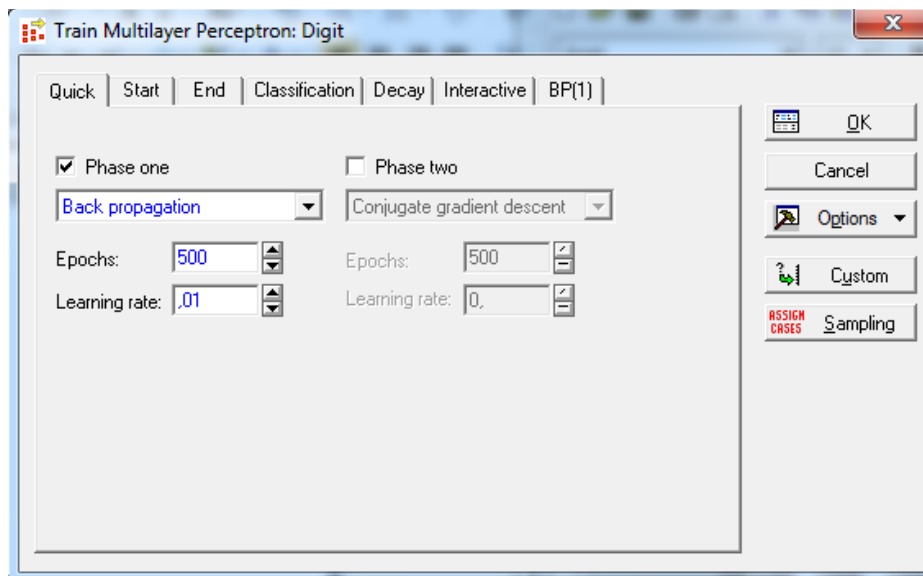


სადაც *Quick* (სწრაფი) ჩანართში შევარჩიეთ ქსელის ტიპი *Multilayer Perceptron* (მრავლობითი პერსეპტრონი).

ამავე ფანჯარაში ჩავრთოდ *Units* (ქსელის ელემენტები) ჩანართი და *Hidden layer 1* ველში მიუთითოთ ფარული შრის ელემენტების რაოდენობა 5.



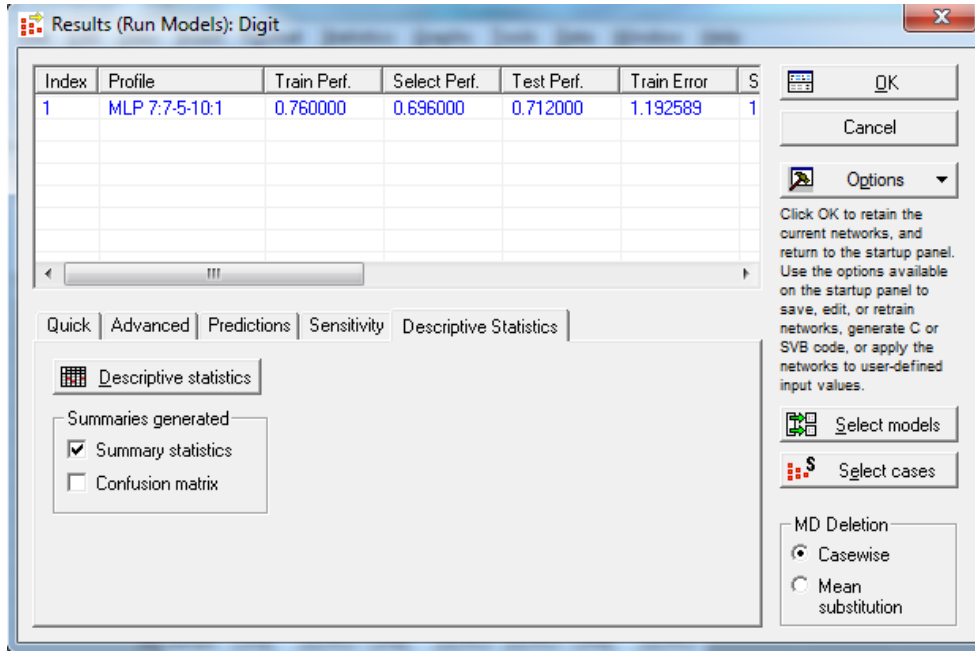
ამდენად, ქსელების არქიტექტურა დადგენილია. *Click OK*. ეკრანზე გამოდის **Train Multilayer Perceptron** (მრავალშრიანი პერსეპტრონის სწავლება)-ის ფაჯარა,



სადაც ჩავრთოდ **Phase one** (პირველი ფაზა) ოპცია და შევარჩიოთ სწავლების **Back propagation** (უკუგავრცელების) ცდომილების უკუგავრცელების ალგორითმი. ეს ალგორითმი შეიძლება გამოყენებულ იყოს **Statistica Neural Networks** პაკეტის ქსელების უმრავლესობისათვის, მაგრამ ის განსაკუთრებით მოსახერხებელია მრავალშრიანი პერსეპტრონების სწავლებისათვის. სხვა სწავლების ალგორითმებთან შედარებით მას სჭირდება ნაკლები მესხიერება და ჩვეულებრივ საკმაოდ სწრაფად აღწევს დასაშვებ მინიმალური ცდომილების დონემდე. **Epochs** (ეპოქები)-ის ველში ჩავწეროთ 500. ეპოქის ქვეშ იგულისხმება სწავლის პროცესში ინტერაციის მაქსიმალური რაოდენობა. ჩვეულებრივად ეპოქების

რაოდენობა იცვლება 100 -დან 1000-მდე. **Learning rate** (სწავლების სიჩქარე) ველში გაჩუმების პრინციპით დაყენებული მნიშვნელობა დავტოვოთ ხელუხლებლად, რადგან ეს სიდიდე უმეტეს ამოცანებისათვის უზრუნველყოფს სწავლების პროცესის მდგრადობას.

უნდა აღინიშნოს, რომ სწავლების პროცედურა უნდა გავიმეოროთ მანამ, სანამ არ მივიღებთ სწავლების მინიმალურ ცდომილებას. სწავლების პარამეტრების დაყენების შემდეგ **Click OK**. ჩაირთვება სწავლების პროცედურა და ეკრანზე გამოდის **Results (Run Models)** შედეგების ფანჯარა,

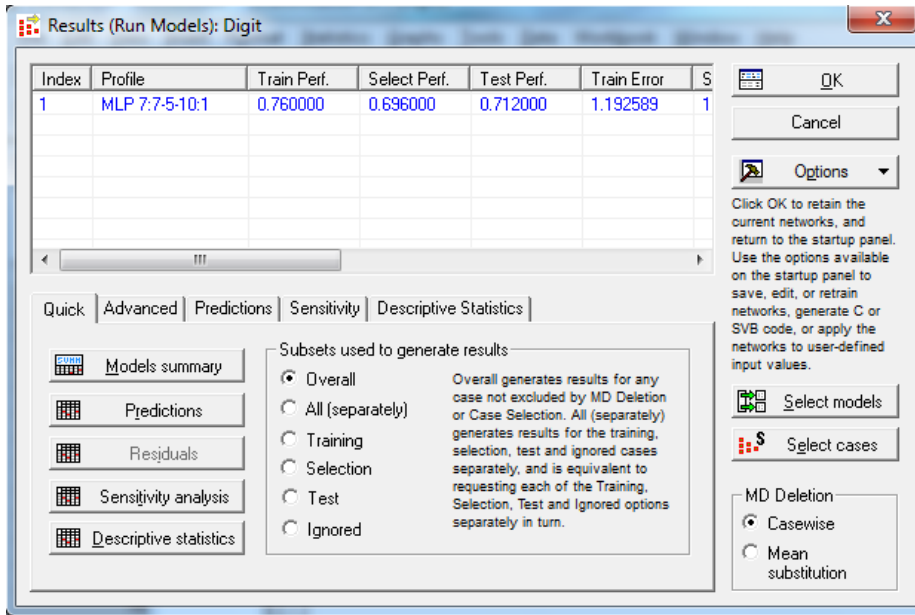


სადაც ჩავრთოდ **Quick** ჩანართი და **Click Descriptive statistics** (აღწერითი სტატისტიკა)-ის ღილაკზე. ეკრანზე გამოდის შეცდომების ცხრილი: **Classification** (კლასიფიკაცია).

	Classification (1) (Digit)									
	DIGIT.seven.1	DIGIT.one.1	DIGIT.four.1	DIGIT.two.1	DIGIT.eight.1	DIGIT.five.1	DIGIT.six.1	DIGIT.zero.1	DIGIT.nine.1	DIGIT.three.1
Total	48,00000	43,00000	52,00000	43,00000	58,00000	54,00000	50,00000	42,00000	55,00000	55,00000
Correct	32,00000	35,00000	43,00000	32,00000	39,00000	40,00000	43,00000	33,00000	33,00000	36,00000
Wrong	16,00000	8,00000	9,00000	11,00000	19,00000	14,00000	7,00000	9,00000	22,00000	19,00000
Unknown	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Correct(%)	66,66667	81,39535	82,69231	74,41860	67,24138	74,07407	86,00000	78,57143	60,00000	65,45455
Wrong(%)	33,33333	18,60465	17,30769	25,58140	32,75862	25,92593	14,00000	21,42857	40,00000	34,54545
Unknown(%)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000

სადაც ათივე კლასისათვის წარმოდგენილია დაწვრილებით ინფორმაცია სწორი და არასწორი კლასიფიცირებული დაკვირვების რაოდენობაზე, კერძოდ **Total** (სულ), **Correct** (სწორი), **Wrong** (მცდარი), **Unknown** (უცნობია) შედეგები და მათი პროცენტული მნიშვნელობები: **Correct (%)**, **Wrong (%)**, **Unknown (%)**.

როგორც ცხრილიდან ჩანს, ყველაზე დიდი შეცდომა დაშვებული იყო „9“ ციფრის გარჩევისას (**Wrong (%) = 40%**).



პრედიქტორების მიმართ მგრძობიარობის ანალიზის ჩასატარებლად *Click Sensitivity Analysis* (მგრძობიარობის ანალიზი)-ის ღილაკზე. ეკრანზე გამოდის ამავე დასახელების ცხრილი,

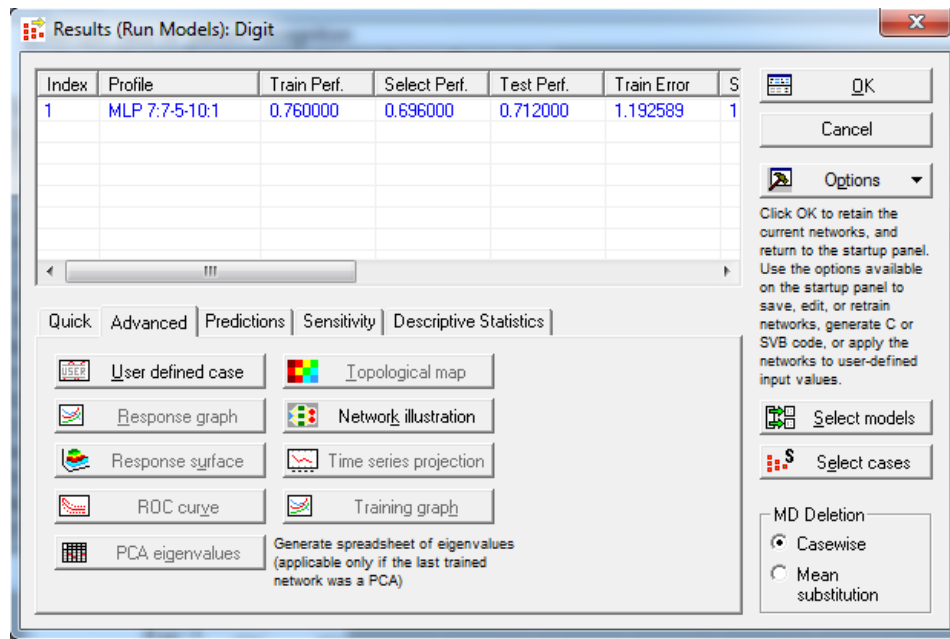
Sensitivity Analysis - 1 (Digit)							
	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7
Ratio.1	1,181192	1,157405	1,318093	1,226775	1,186577	1,075947	1,125123
Rank.1	4,000000	5,000000	1,000000	2,000000	3,000000	7,000000	6,000000

სადაც შვიდივე პრედიქტორისთვის წარმოდგენილია **Ratio** (ფარდობა) და **Rank** (რანგი) მნიშვნელობები. **Rank.1** შეესაბამება უდიდეს მგრძობიარობას მოცემული პრედიქტორისათვის. როგორც ცხრილიდან ჩანს, **VAR3** (ზედა მარჯვენა ვერტიკალური) და **VAR4** (შუა ჰორიზონტალური) ცვლადებს შეაქვთ დიდი წვლილი ციფრების გარჩევაში.

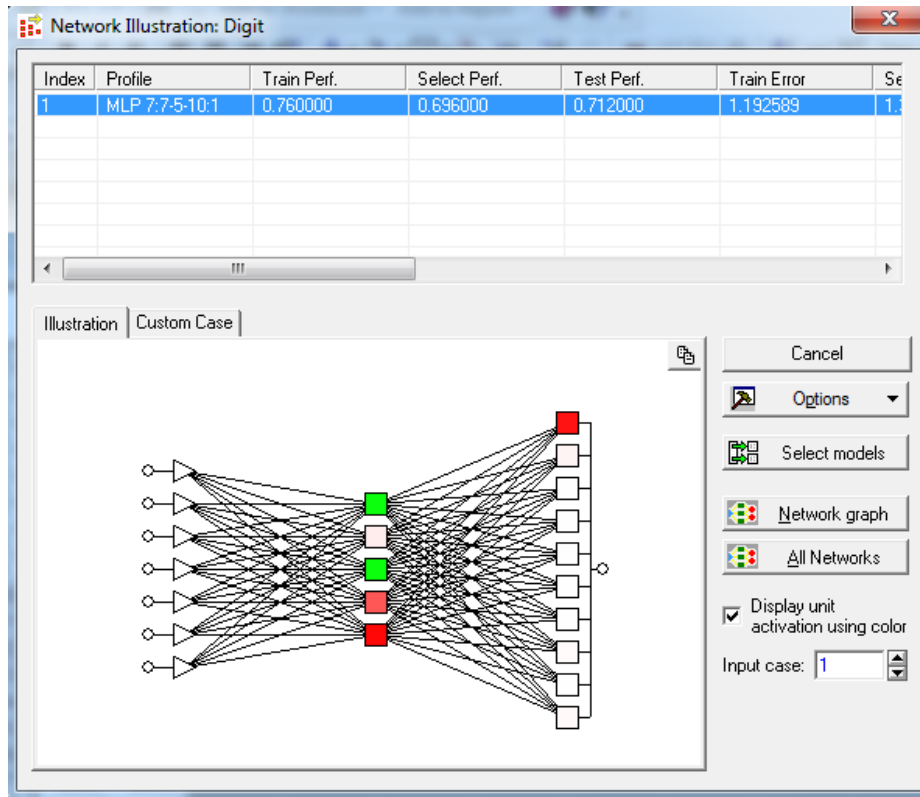
მოდელის შედეგი სასწავლო, საკონტროლო, და ტესტური სიმრავლეებისათვის შეიძლება მივიღოთ **Predictions** (პროგნოზირება) ღილაკზე დაჭერით. პროგნოზირების ცხრილი პირველი 15 დაკვირვებისთვის ნაჩვენებია შემდეგ ნახაზზე:

	Prediction (1) (Digit)	
	DIGIT	DIGIT.1
1	seven	seven
2	one	one
3	four	four
4	two	eight
5	eight	eight
6	one	one
7	five	five
8	six	six
9	two	two
10	eight	nine
11	one	one
12	eight	eight
13	one	seven
14	seven	seven
15	seven	one

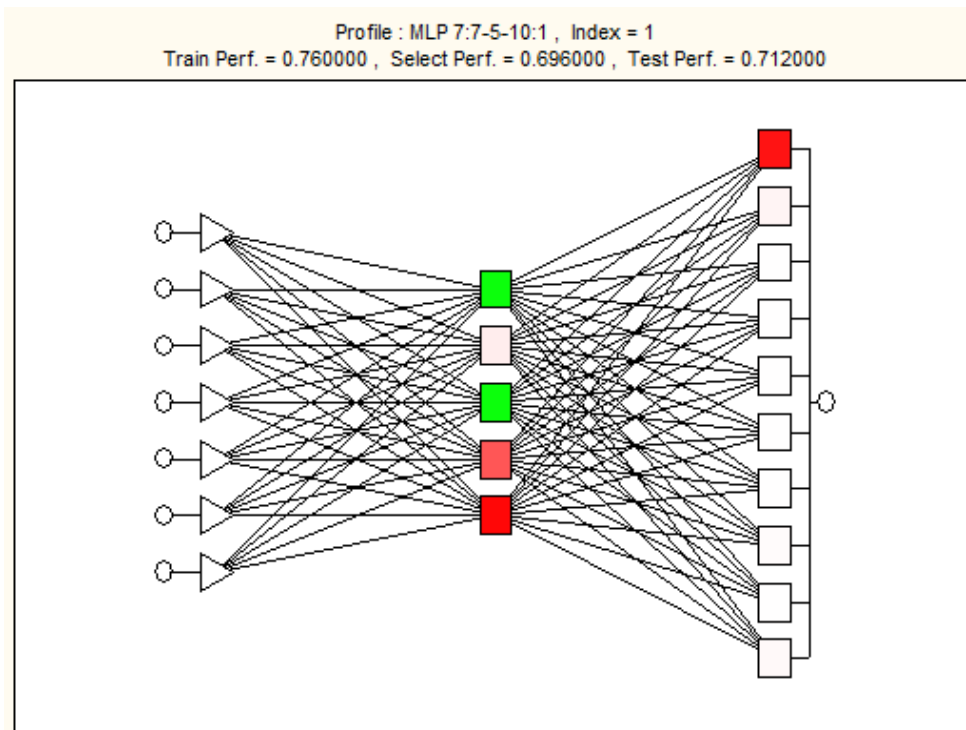
აგებული ქსელის არქიტექტურა მოსახერხებელია წარმოვანით გრაფიკული სახით. გრაფიკის ასაგებად **Results (Run Models)** შედეგების ფანჯარაში ჩავრთოდ **Advanced** ჩანართი



და **Click Network illustration** (ქსელის ილუსტრაცია)-ის ღილაკზე. ეკრანზე გამოდის ამავე დასახელების ფანჯარა.



თუ გვინდა ნეირონული ქსელის ცალკე გრაფის სახით წარმოდგენა, მაშინ *Click Network graph* ღილაკზე



ქსელის აქტივაციის დონე აღნიშნულია ფერებით: წითელი – დადებითი აქტივიზაციისათვის, მწვანე – უარყოფითი აქტივიზაციისათვის.

პრაქტიკული სამუშაო 4

პროგნოზირების ამოცანის გადაწყვეტა

მაგალითისათვის განვიხილოთ სრული სამი წლის ნავთობის ფასის მონაცემები. ამასთანავე ნავთობის ფასის ცვლა დაფიქსირებულია ყოველ სავაჭრო დღეს. უნდა გაავანალიზოთ დროითი მწკრივი (*Time Series*) და გამოვეყოთ ნავთობზე ფასის მერყეობის დამახასიათებელი პერიოდები. დროითი მწკრივის მოდელის დახმარებით საჭიროა ავაგოთ ნავთობის გაყიდვის პროგნოზი წინსვლით რამდენიმე თვით ადრე და შევაფასოთ პროგნოზის სიზუსტე.

Statistica სისტემის *Datasets* ბიბლიოთეკიდან გავხსნათ *OIL* ნავთობის ფასის მონაცემთა ფაილი, რომლის ფრაგმენტი წარმოდგენილია შემდეგ ნახაზზე:

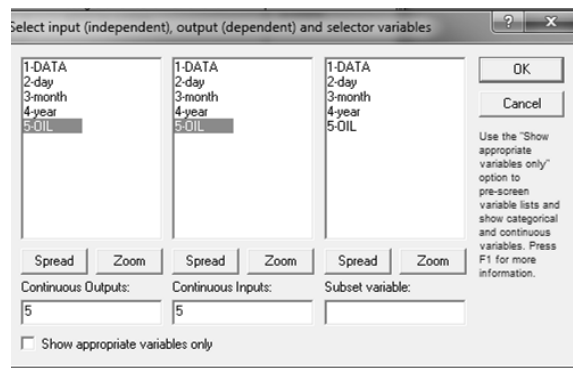
	OIL Price				
	1	2	3	4	5
	DATA	day	month	year	OIL
1	27/12/95	27	12	95	18.73
2	28/12/95	28	12	95	18.78
3	29/12/95	29	12	95	18.65
4	02/01/96	2	1	96	18.95
5	03/01/96	3	1	96	19.15
6	04/01/96	4	1	96	19.13
7	05/01/96	5	1	96	19.50
8	08/01/96	8	1	96	19.68
9	09/01/96	9	1	96	19.05
10	10/01/96	10	1	96	18.48
11	11/01/96	11	1	96	17.85
12	12/01/96	12	1	96	17.58
13	15/01/96	15	1	96	17.50
14	16/01/96	16	1	96	17.73
15	17/01/96	17	1	96	17.43
16	18/01/96	18	1	96	17.88
17	19/01/96	19	1	96	17.53
18	22/01/96	22	1	96	17.53
19	23/01/96	23	1	96	17.38
20	24/01/96	24	1	96	17.33
21	25/01/96	25	1	96	17.00
22	26/01/96	26	1	96	16.70
23	29/01/96	29	1	96	16.23
24	30/01/96	30	1	96	16.53
25	31/01/96	31	1	96	16.63
26	01/02/96	1	2	96	16.78
27	02/02/96	2	2	96	17.05

სადაც **DATA**(თარიღი) ცვლაში ჩაწერილია თარიღები, ხოლო ცვლადები 2–4 წარმოადგენენ თარიღის ჩანაწერს სამი ცვლადის საშუალებით **day, month, year**, (დღე/ თვე /წელი).

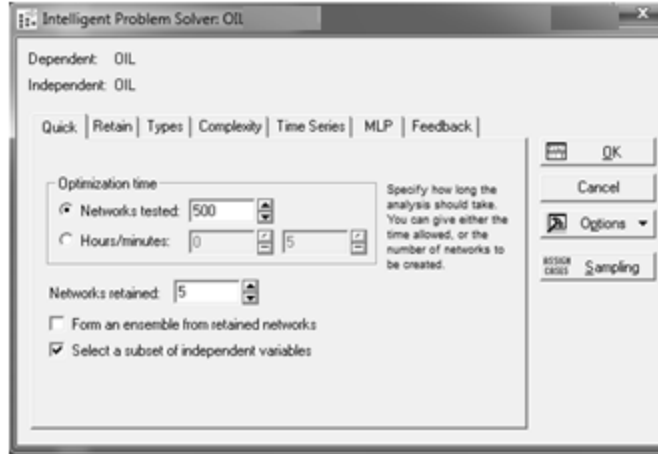
სასტარტო პანელის **STATISTICA Neural Networks** გამოსაძახებლად, მენიუდან **Statistics** შევარჩიოთ ბრძანება **Neural Networks**. ეკრანზე გამოდის **STATISTICA Neural Networks(SNN)** ფანჯარა.



ჩაერთოდ **Quick** ჩანართი და **Problem type** ველში შევარჩიოთ ამოცანის ტიპი **Time Series** (დროითი მწკრივები), ხოლო **Select analysis** ველში ამოხსნის მეთოდი – **Intelligent Problem Solver** (გადაწყვეტის ოსტატი). ცვლადების მოსანიშნად **Click Variables** ლიკაკზე და ეკრანზე გამოსულ **Select input(independent), outp(dependent) and Selector vaeiables** (მოვნიშნოთ შემავალი (დამოკიდებული), გამოშვებული (დამოკიდებული) და დამაჯგუფებელი ცვლადები) მოვნიშნოთ **OIL**.

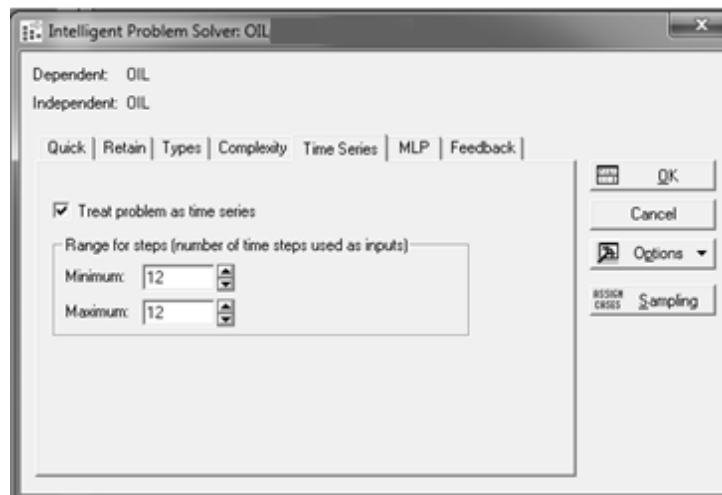


შემდეგ **Click OK**. ვბრუნდებით **STATISTICA Neural Networks (SNN)** ფანჯარაში და კვლავ **OK**. ეკრანზე გამოდის **Intelligent Problem Solver** ფანჯარა.

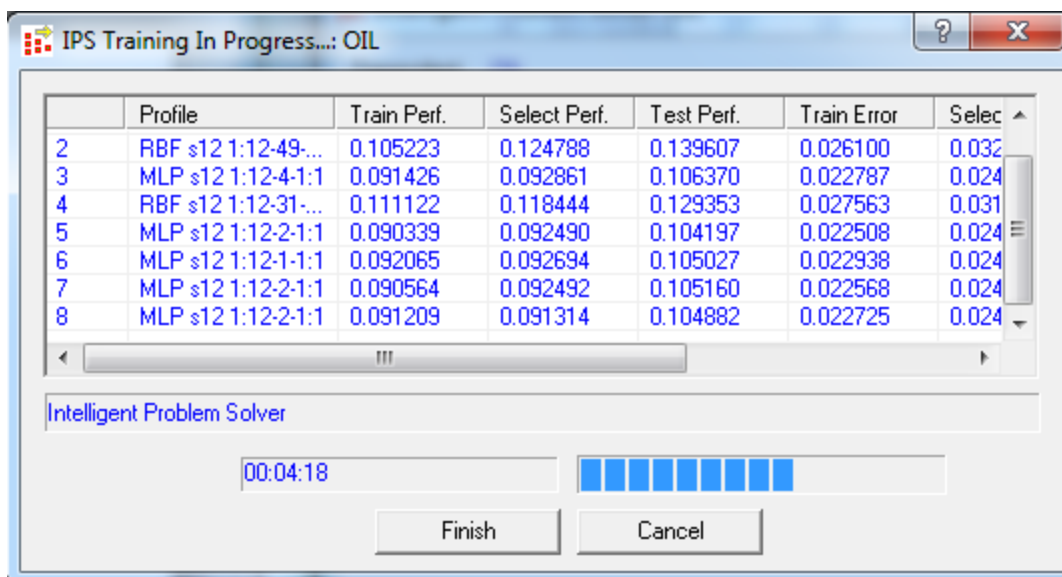


ჩავრთოდ **Quick** (სწრაფი) ჩანართი. ჯგუფში **Optimization time** (ანალიზის ხანგრძლივობა) ველში **Networks tested** (ქსელების ტესტირება) ჩავწეროთ 500. მოცემული პარამეტრი განსაზღვრავს ჩატარებული ანალიზის ხანგრძლივობას (ტესტირებისთვის განკუთვნილი მოდელების რიცხვი). ქსელის შესანახად **Networks retained** (ქსელის შენარჩუნება) ველში ჩავწეროთ ციფრი 5, რათა **Intelligent Problem Solver** (გადაწყვეტის ოსტატი) შეინახოს 5 საუკეთესო ქსელის ვარიანტი. **Types** ჩანართით შევარჩიონ ქსელის ტიპი

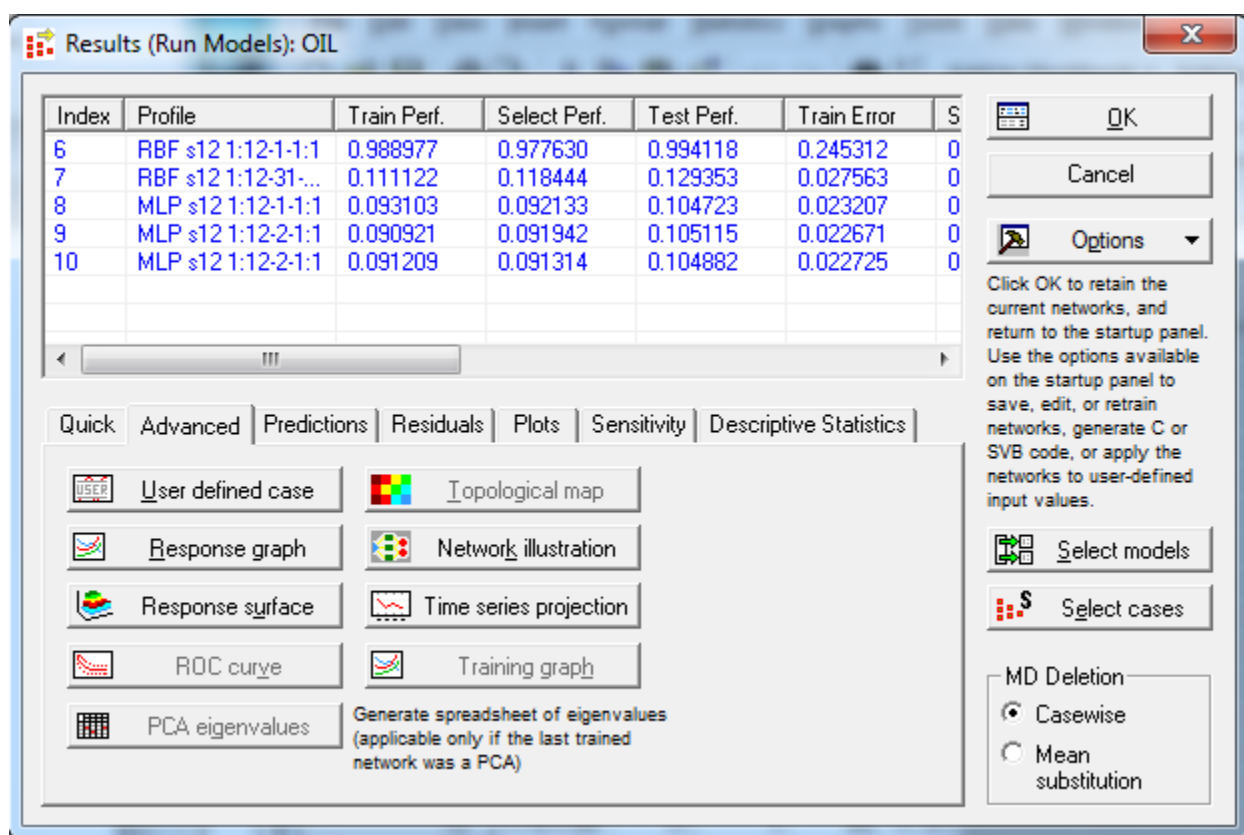
ჩავრთოდ **Time Series** ჩანართი და მოვნიშნოთ ოპცია **Treat problem as time series** (მონაცემთა ინტერპრეტაცია როგორც დროის მწკრივი) .



პროგნოზირების ფანჯრად ავიღოთ 12, ამიტომ **Range for steps (number of steps used as inputs)** ჩარჩოს მინიმუმ და მაქსიმუმ ველებში ჩავწეროთ 12, რომელიც შეესაბამება იმას, რომ ერთი წლის მონაცემებით მოხდება მომდევნო ერთი თვის მნიშვნელობის პროგნოზირება. **Click OK**. დაიწყება მოდელის შესწავლა.



მოდელის სწავლების შემდეგ ეკრანზე გამოჩნდება **Results** (შედეგები) ფანჯარა

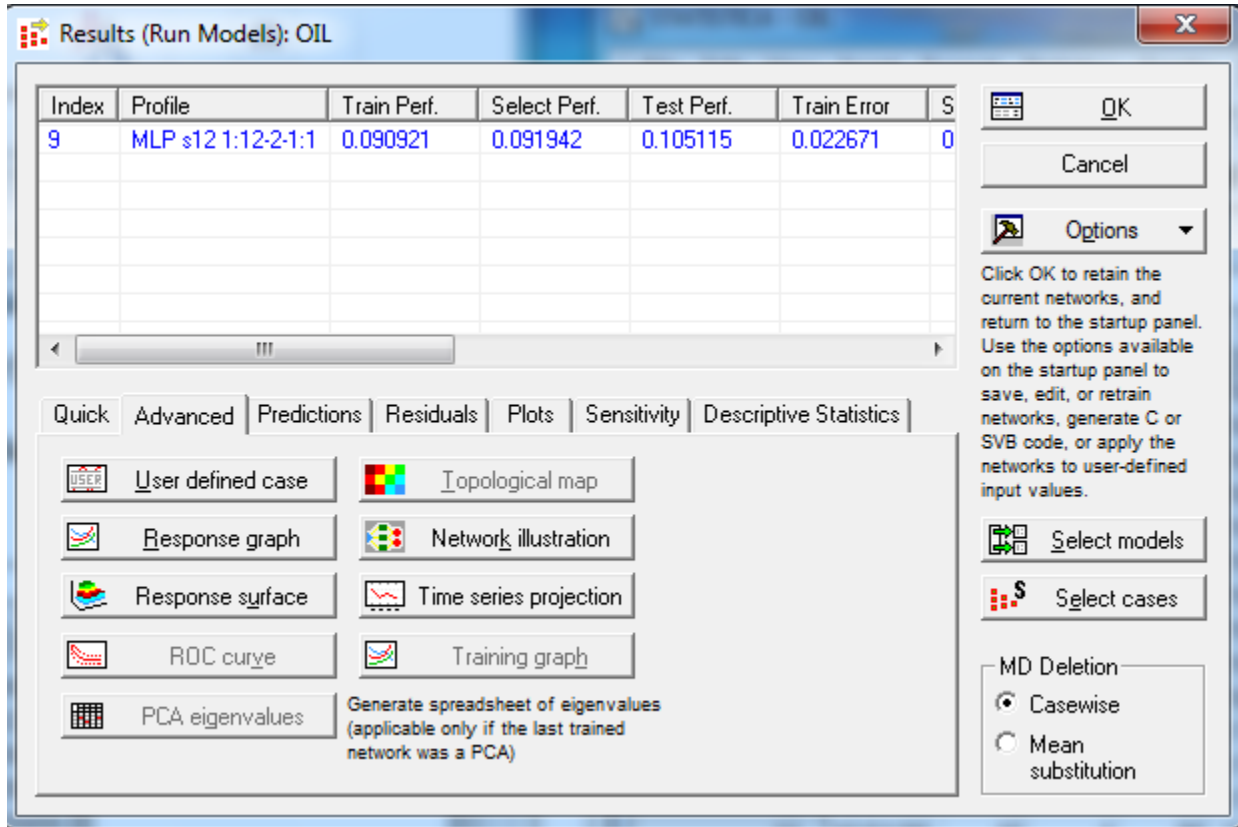


ცხრილის ზედა ნაწილში იმყოფება შენახული ქსელების არქიტექტურა და აგებული მოდელის სხვადასხვა ხარისხის მაჩვენებელი.

გავხსნათ *Dsscriptive Statistics* ჩანართი და *Click Dsscriptive Statistics* ღილაკზე. ეკრანზე გამოდის შედეგების ცხრილი

	Regression (6-10) (OIL)				
	OIL.6	OIL.7	OIL.8	OIL.9	OIL.10
Data Mean	17,27094	17,27094	17,27094	17,27094	17,27094
Data S.D.	4,06047	4,06047	4,06047	4,06047	4,06047
Error Mean	-0,04965	-0,02728	-0,01291	-0,01033	-0,02783
Error S.D.	4,00649	0,47723	0,38842	0,38443	0,38438
Abs E. Mean	3,35501	0,37364	0,29300	0,28813	0,28824
S.D. Ratio	0,98670	0,11753	0,09566	0,09468	0,09466
Correlation	0,16328	0,99307	0,99542	0,99551	0,99551

სადაც *SD Ratio* მაჩვენებლით და სწავლების ცდომილებით შევარჩიოთ მე-9 მოდელი



დაკვირვების გრაფიკის ასაგებად ჩავრთოდ ჩანართ *Plots* (გრაფიკი) და *X-axis* (ღერძზე)-ის ველში შევარჩიოთ *Predicted* (პროგნოზირება), ხოლო *Y-axis* (ღერძზე) – *Observed* (დაკვირვება).

Results (Run Models): OIL

Index	Profile	Train Perf.	Select Perf.	Test Perf.	Train Error	S
9	MLP s12 1:12-2-1:1	0.090921	0.091942	0.105115	0.022671	0

Quick | Advanced | Predictions | Residuals | **Plots** | Sensitivity | Descriptive Statistics

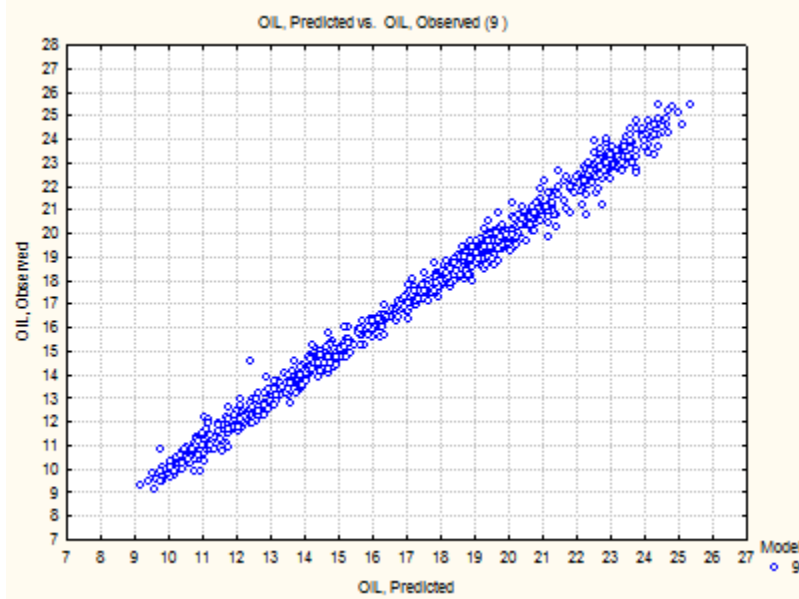
X-axis: Predicted
Y-axis: Observed
Variable: OIL

Graph X versus Y
Histogram of Y

Select models
Select cases

MD Deletion
 Casewise
 Mean substitution

აგეთ X და Y ცვლადების დამოკიდებულების გრაფიკი. ამისათვის
 Click **Graph X versus Y** ღილაკზე. ეკრანზე გამოდის დამოკიდებულების გრაფიკი.



საიდანაც ჩანს, რომ გრაფიკი ფაქტიურად იდეალურად დევს სწორ ხაზზე. ამ შედეგის მიღება მოსალოდნელი იყო, რადგან როგორც შედეგების **Results (Run Models)** ფანჯარაში ჩანს მე-9 მოდელის სწავლების ცდომილება შეადგენს მხოლოდ 2,27%, ხოლო ტესტირების ცდომილება – 2,5%.

შედეგების **Results(Run Models)** ფანჯარის **Quick** ჩანართში **Click Descriptive statistics** (აღწერითისტატისტიკა)-ის ღილაკზე. ეკრანზე გამოდის რეგრესიის შეცდომების ცხრილი,

	Regression (9) (OIL)	
	OIL.9	
Data Mean	17,27094	
Data S.D.	4,06047	
Error Mean	-0,01033	
Error S.D.	0,38443	
Abs E. Mean	0,28813	
S.D. Ratio	0,09468	
Correlation	0,99551	

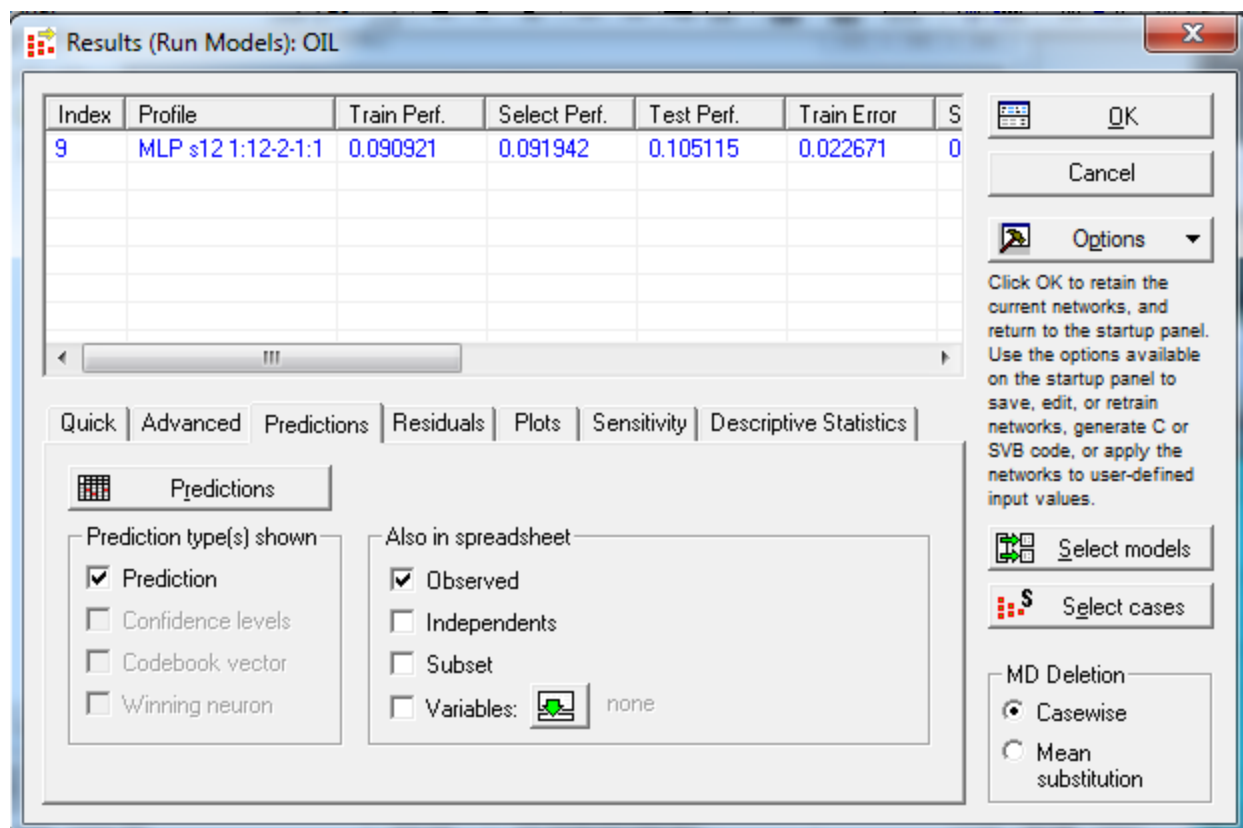
მიღებულ შედეგებიდან ჩანს, რომ მე-9 ქსელის სტანდარტული გადახრების ფარდობის სიდიდე **S.D.Ratio** 0,09 ტოლია, რაც საკმაოდ კარგი მაჩვენებელია.

თუ შედეგების **Results(Run Models)** ფანჯარის **Subets used to generale resultsc** (შერჩევის შედეგების გამოყვანა)-ის ველში **Overall** (სრული) ოპციის მაგივრად ჩავრთავთ **All (separately)** (ყველა ცალკე) ოპციას, მაშინ ეს მოგვცემს საშუალებას მივიღოთ ცხრილი, სადაც შერჩევის შედეგები ყველა ამონარჩევისათვის (სასწავლო, საკონტროლო, ტესტური და იგნორირებული) ცალკე გამოდის.

	Regression (9) (OIL)			
	T.OIL.9	S.OIL.9	X.OIL.9	I.OIL.9
Data Mean	17,23265	16,97906	17,64076	0,00
Data S.D.	4,02374	4,26405	3,89514	0,00
Error Mean	0,01482	-0,03275	-0,03910	0,00
Error S.D.	0,36584	0,39205	0,40944	0,00
Abs E. Mean	0,27919	0,29031	0,30415	0,00
S.D. Ratio	0,09092	0,09194	0,10511	0,00
Correlation	0,99586	0,99578	0,99451	0,00

როგორც ცხრილიდან ჩანს **S.D.Ratio** სიდიდე ყველა ამონარჩევისათვის თითქმის ერთნაირია.

იმისათვის, რომ საბოლოოდ დავრწმუნდეთ ნეიროქსელის შერჩევის სისწორეში **Results(Run Models)** ფანჯარაში ჩავრთოდ **Predictions** ჩანართი



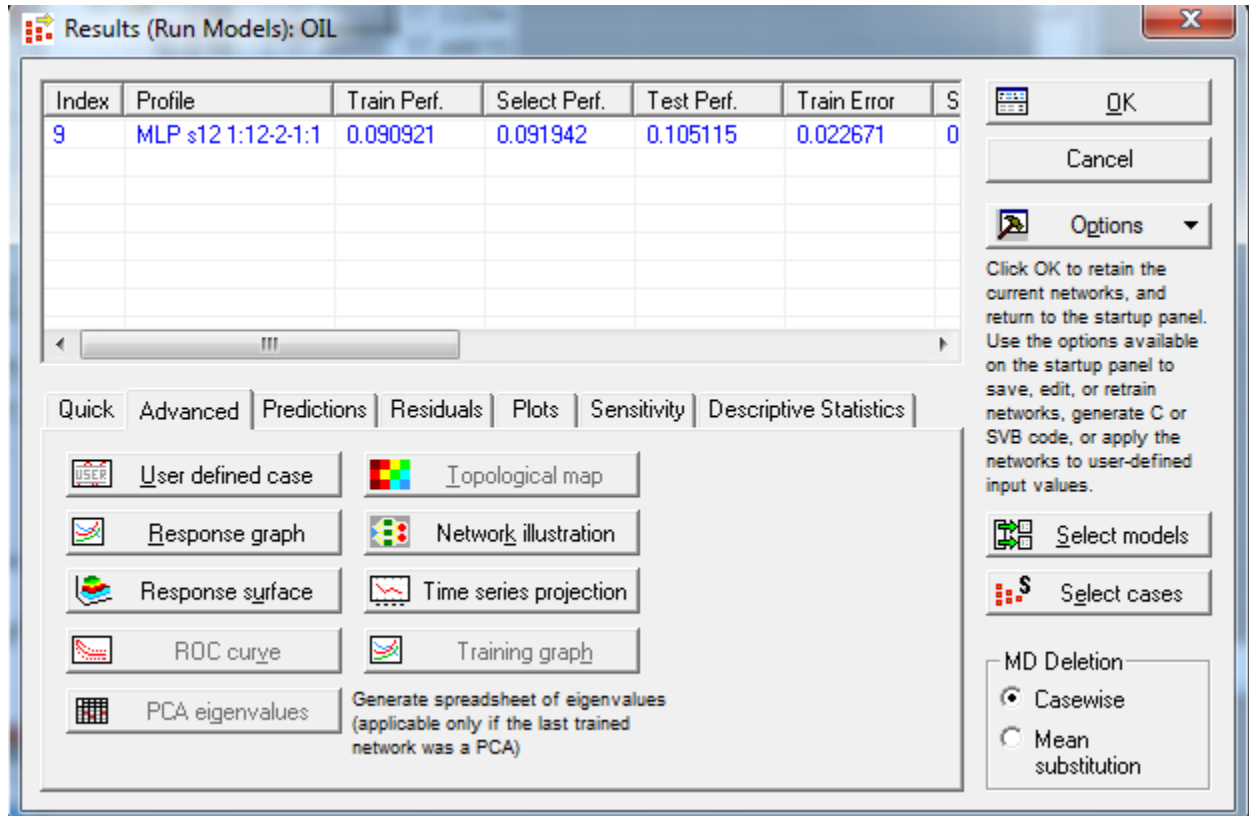
და **Click Predictions** (პროგნოზირება)-ის დილაკზე. ეკრანზე გამოდის შედეგების ცხრილი,

	Prediction (9) (Test) (OIL)	
	OIL	OIL.9
14	17,73000	17,53294
19	17,38000	17,44610
25	16,63000	16,52752
32	17,23000	17,01658
36	18,18000	18,29873
41	19,28000	19,06372
43	18,85000	18,96416
46	18,80000	18,66067
50	19,18000	18,74652
52	18,85000	19,08645
61	20,35000	20,40001
64	21,68000	21,33770
65	21,33000	21,66931
66	20,25000	21,42842
68	20,80000	20,31212
69	21,18000	20,78403
75	22,73000	23,78474

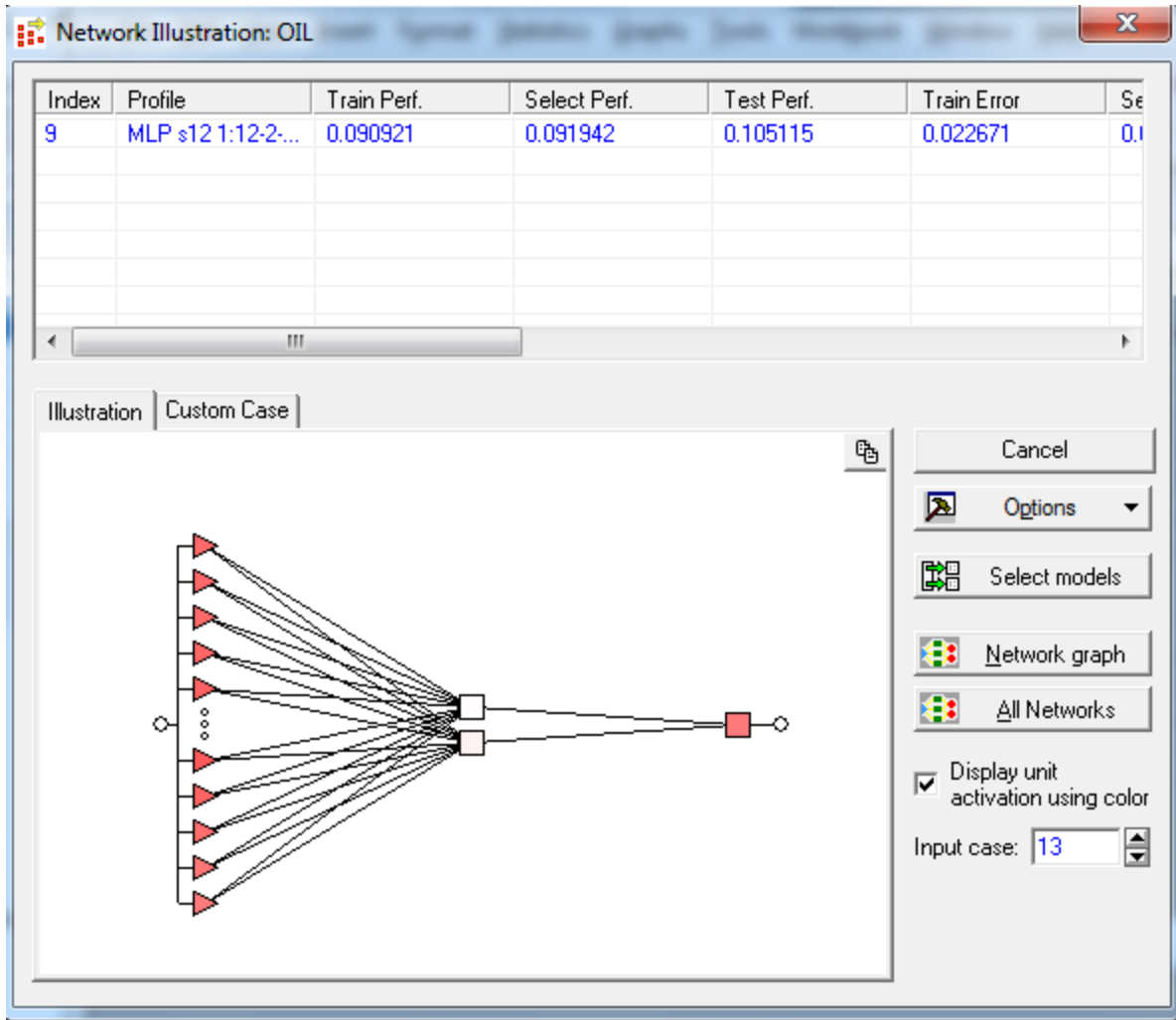
სადაც მარცხენა სვეტში მოცემულია **OIL**-ის ჯეშმარიტი მნიშვნელობები, ხოლო მარჯვენა სვეტში – ნეიროქსელის მიერ ტესტური ამონარჩევით მიღებული

პროგნოზირებული მნიშვნელობები. როგორც ცხრილიდან ჩანს, მიღებული შედეგები დამაკმაყოფილებელია. ნეიროქსელის შერჩევა შესაძლებელია აგრეთვე, კროს-შემოწმების მეთოდით (იხ. პრაქტიკულსამუშაო 1).

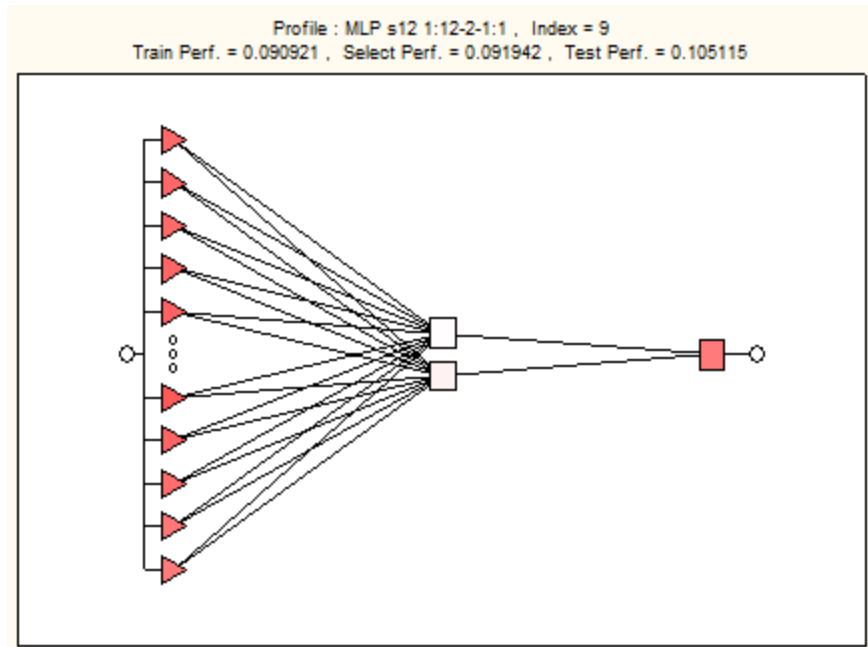
ნეირონული ქსელის არქიტექტურის ასაგებად შედეგების **Results(Run Models)** ფანჯარაში ჩავრთოდ **Advanced** ჩანართი



და **Click Network illustration** (ქსელის სქემა)-ის ლილაკზე. ეკრანზე გამოდის ამავე დასახელების ფანჯარა.



თუ გვინდა ნეირონული ქსელის ცალკე გრაფის სახით წარმოდგენა, მაშინ *Click Network graph* ღილაკზე



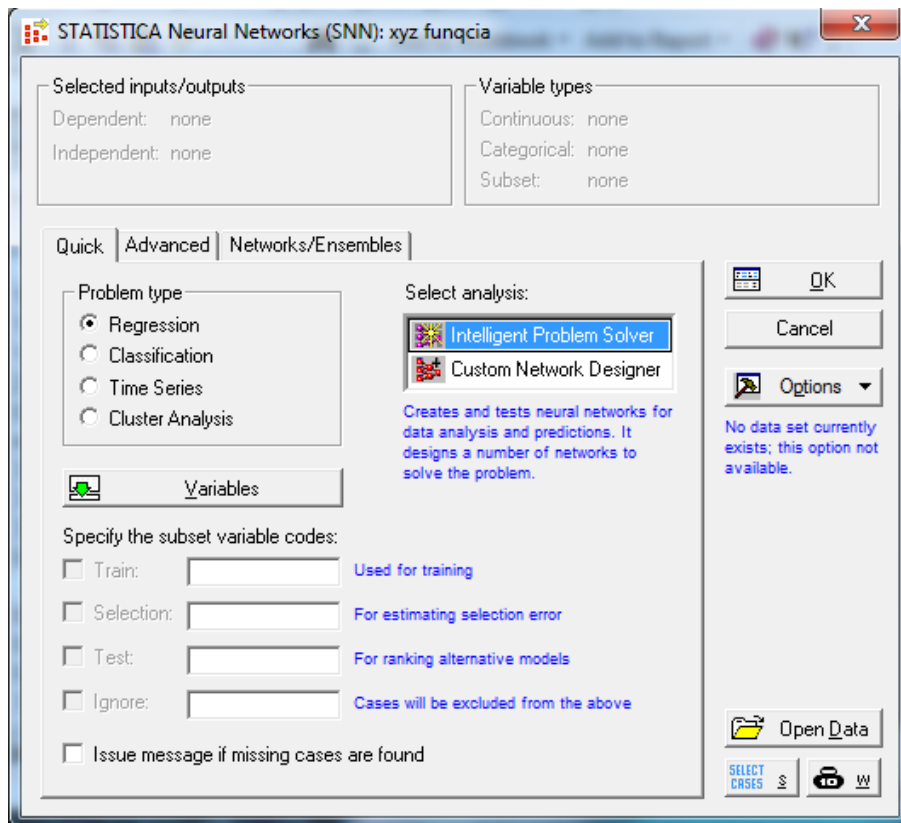
პრაქტიკული სამუშაო 5

ცვლადების შერჩევა და განზომილების შემცირება

პროგრამის მუშაობა განვიხილოთ *pd 2* მოდელური ფაილისათვის, რომელსაც გააჩნია დამოკიდებული Y ცვლადი და რამოდენიმე ათეული დამოუკიდებელი ცვლადები (პრედიქტორები). უნდა აღინიშნოს, რომ ფაილი შეიცავს პრედიქტორების დიდ რაოდენობას, კერძოდ 45, ამიტომ მიზანშეწონილია შევარჩიოთ ინფორმატიული ცვლადები და ამით მოვახდინოთ განზომილების შემცირება. საწყისი მონაცემების ფაილის ფრაგმენტი წარმოდგენილია შემდეგ ნახაზზე:

	1 F1	2 F2	3 F3	4 F4	5 F5	6 F6	7 F7	8 F8	9 F9	10 F10	11 F11	12 F12	13 F13	14 F14	15 F15
1	1024	2435	711,5	16,055	16,94	13,23555	0,6209	98,6549	10,2	14	108,6499	85,9	78,6	8,09	6,74
2	1011	2459,6	774,6	8,625	8,855	12,61	1,69576	100,394	90	50	111,9129	99,3	62,5	8,34	14,14
3	1117	2599	804,4	21,865	22,915	16,1	1,29677	98,8863	12	12,2	96,23004	119,8	58,8	4,77	36,36
4	945	2367	533	28,05	29,93	15,15548	2,74386	99,1474	86	10,21	95,68411	115,5	36,2	4,78	33,8
5	959	2376	350	25,34	25,785	19,65147	2,30093	102,662	11,7	94,94796	111,5913	118,3	78,8	3,07	10,22
6	1008	2436	855,5	10,355	10,51	18,56315	2,07885	104,996	8	98,45073	94,20351	115,4	82,1	122	15,68
7	1025,2	2449	550	30,21	32,375	18,02188	3,0341	97,9636	19,8	95,60386	109,416	184,9	64,7	96	170
8	1023	2439,6	350	5,705	5,745	21,36781	3,16542	100,687	22,94	91,62065	108,9486	161,7	73,8	1,02	4,42
9	1006	2424	844,4	18,15	18,455	21,69199	3,99269	99,7761	10,3	93,51479	107,1998	127,6	51	3,24	11,4
10	1013,2	2427	750	66,07	72,395	23,71771	3,165	97,6319	49	94,31208	113,5048	166,8	82,2	4,76	3,7
11	996	2396	533	40,42	43,485	21,69096	13,23555	103,168	13,2	96,86733	94,77887	158,6	49,5	540	11,98
12	1047	2456	650,4	4,955	5,11	22,60129	12,61308	98,0758	12,89	94,79189	100,0957	138,6	67	629	31,84
13	1109	2528	850	27,63	29,42	25,41543	16,0987	96,3053	13,2	15,5	110,7379	88,9	859	42,8	2,66
14	1016	2441	657	30,035	33,22	26,91658	15,15548	96,7556	12,7	93,41488	96,55905	78,4	612	76,2	6,2
15	1061	2495	450	3,72	4,11	25,56655	19,65147	100,778	13,9	96,22288	107,925	42	451	69,5	10,7
16	940	2364	362,3	0,895	0,92	25,00305	18,56315	56,3381	10,9	36,3	112,379	142	70,8	364	118
17	1149	2565	459,9	50,935	52,57	24,91046	18,02188	52,2928	10,2	49,7	103,6192	94,9	53	209	3,32
18	1026	2326	450	24,16	16,47	26,47922	7,03438	51,8912	7,7	62,5	112,7804	110,4	55,2	11,4	8,84
19	935	2539	362,3	34,41	18,66	29,61682	9,380429	51,3659	15,4	6,7	90,07166	92,8	54,9	378	2,92
20	1062,8	2426	459,9	26,82	14,09	29,9272	10,48961	65	16,9	7,1	93,00613	167,2	698	61,1	6,34
21	947	2356	606	26,42	12,27	28,60062	4,1136	23,2	96	4,78	112,3285	153,9	1061	56,3	77,1407
22	1682	2594	675	29,29	15,55	28,43999	10,46996	45,3	70	2,75	98,08788	184,3	1008	96,7	78,18
23	903	2412	811,5	11,255	13	28,99903	3,04167	63	16,3	85	107,4324	114,3	816	68	8,4186

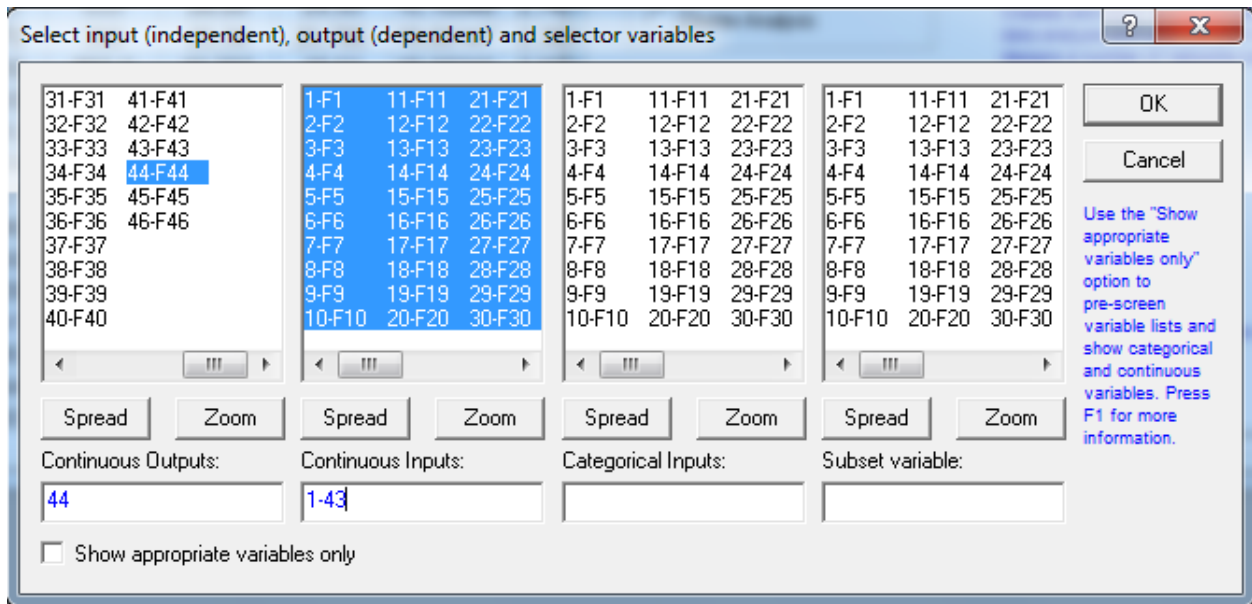
სასტატო პანელის *STATISTICA Neural Networks* გამოსაძახებლად, მენიუდან *Statistics* შევარჩიოთ ბრძანება *Neural Networks*. ეკრანზე გამოდის *STATISTICA Neural Networks(SNN)* ფანჯარა,



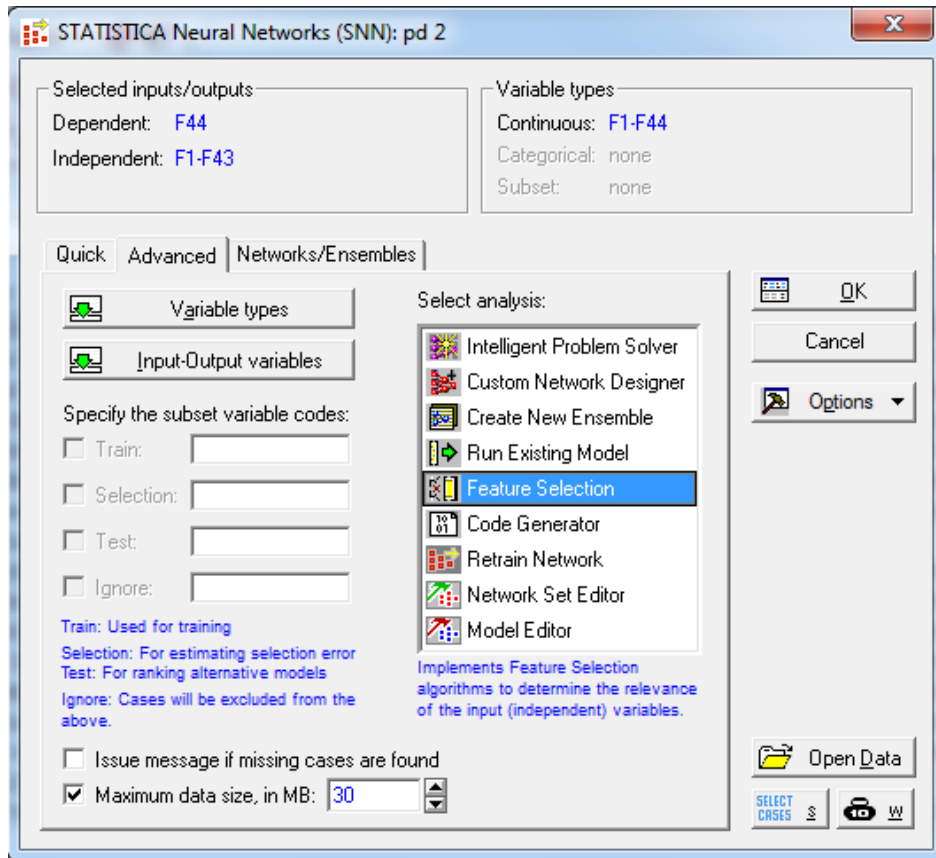
სადაც ჩავრთოდ **Quick** ჩანართი და **Problem type** ველში შევარჩიოთ ამოცანის ტიპი **Regression**, ხოლო **Select analysis** ველში ამოხსნის მეთოდი – **Intelligent Problem Solver** (გადაწყვეტის ოსტატი).

ცვლადების მოსანიშნად **Click Variables** ღილაკზე და ეკრანზე გამოსულ ფანჯარაში ავირჩიოთ ცვლადები. მოცემულ მაგალითში არის გამომაგალი (დამოკიდებული) ცვლადი და 46 შემავალი ცვლადი.

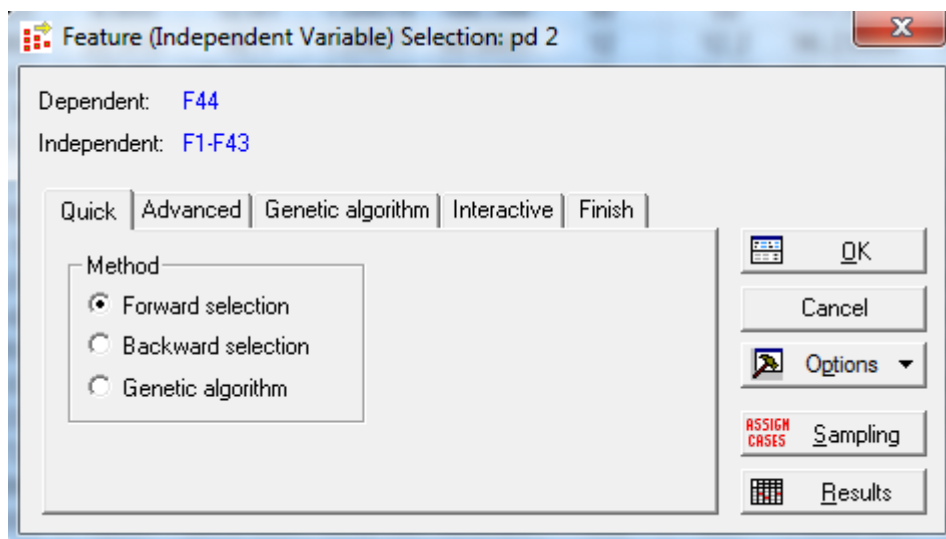
მოვნიშნოთ **Continuous Outputs** (უწყვეტი გამომაგალი) ველში - y ცვლადი, რომლის ნომერია 44, ხოლო **Continuous Inputs** (უწყვეტიშემავალი) ველში – დამოუკიდებელი ცვლადები 1-43. რადგან არ გაგვანია კატეგორიული ცვლადი, ამიტომ(**Categorical Inputs**) ველში ცვლადი არ მოინიშნება. შემდეგ **OK**.



STATISTICA Neural Networks(SNN) ფანჯარაში ჩავრთოდ **Advance** (დამატებითი) ჩანართი და ამოვირჩიოთ ინსტრუმენტი **Feature Selection** (განზომილების შემცირება).



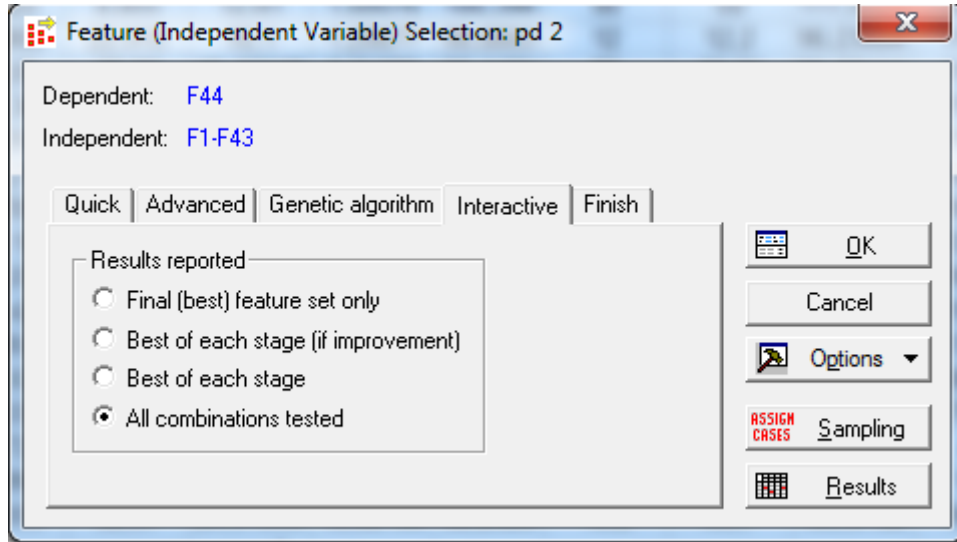
შემდეგ *Click OK* ღილაკზე. ეკრანზე გამოჩნდება განზომილების შემცირების დიალოგის ფანჯარა **Feature (Independent Variable) Selection**,



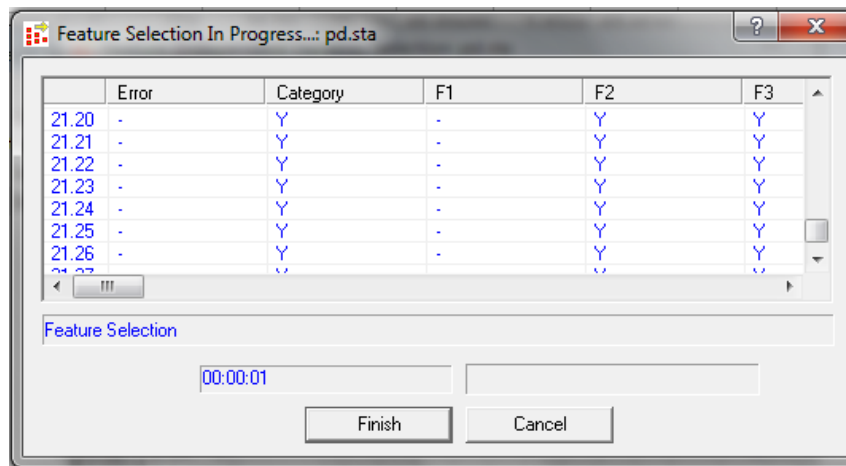
სადაც **Quick** ჩანართში ხელმისაწვდომია შემდეგი მეთოდები: **Forward selection** (თანმიმდევრული ჩართვით), **Backward selection** (თანმიმდევრული გამორიცხვით), **Genetic algorithm** (გენეტიკური ალგორითმი).

1. თანმიმდებრული ჩართვის მეთოდი.

Feature (Independent Variable) Selection ფანჯარაში ჩავრთოდ **Forward selection** (თანმიმდებრული ჩართვით) ოპცია. გავსსნათ **Interactive** (ინტერაქტიური) ჩანართი და ამოვირჩიოთ **All combinations tested** (ტესტირების ყველა კომბინაცია).



შემდეგ *Click OK* ღილაკზე. ეკრანზე გამოჩნდება ნეიროქსელის სწავლების დიალოგიური **Feature Selection In Progress** (მიმდინარეობს შესწავლა) ფანჯარა.



ალგორითმის მუშაობის შედეგი წარმოდგენილია **Forwards Selection** ცხრილში.

Forwards Selection (pd 2)																						
Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22
1.1	0,041686	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.2	0,041665	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.3	0,041718	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.4	0,042087	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.5	0,041817	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.6	0,041690	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.7	0,041736	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.8	0,041597	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-
1.9	0,041953	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-
1.10	0,041654	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-
1.11	0,041683	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-
1.12	0,041624	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-
1.13	0,041662	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-
1.14	0,041662	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-
1.15	0,041686	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
1.16	0,041667	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
1.17	0,041589	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-
1.18	0,041599	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-
1.19	0,041666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-
1.20	0,041673	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-
1.21	0,041673	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y

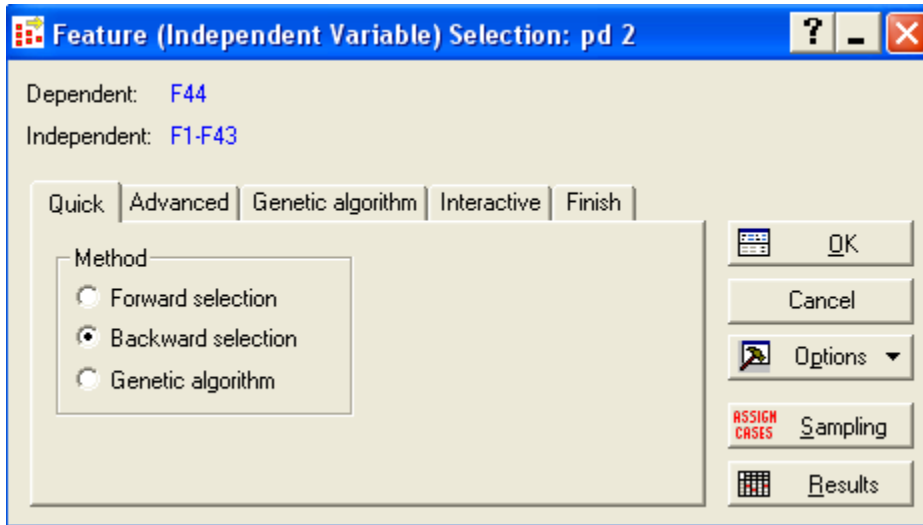
თავდაპირველად ცვლადები ცალ-ცალკე ირთვებიან მოდელში. იძებნება ის პირველი ცვლადი, რომელიც იძლევა ყველაზე მცირე შეცდომას. შემდეგ მოდელი იწყებს მეორე ცვლადის ძებნის პროცედურას, რომელიც ამცირებს შეცდომის მნიშვნელობას და ა.შ. მანამდე სანამ ახალი ცვლადების მოდელში ჩართვა იწვევს ცდომილების შემცირებას. საბოლოო შედეგი ჩაწერილია ცხრილის ბოლო სტრიქონში (*Final*).

Forwards Selection (pd 2)																						
Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22
8.38	0,041009	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	-
8.40	0,041010	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	Y	-	-	-	-
8.41	0,041005	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	Y	-	-	-
8.42	0,041000	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	Y	-	-
8.43	0,041011	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	Y	-
Final	0,040919	-	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	-	-	-	-

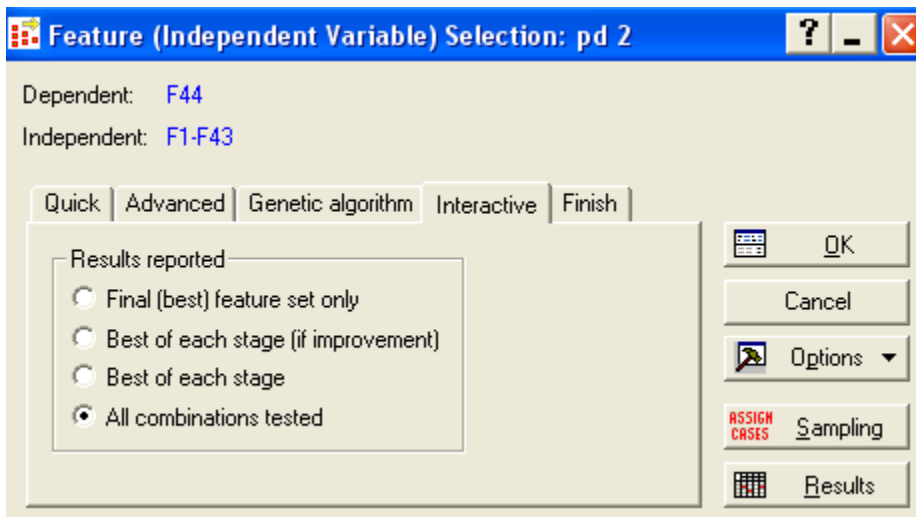
როგორც მიღებული შედეგიდან ჩანს მოდელის მიერ შერჩეული მნიშვნელოვანი ცვლადები შემდეგია: F8, F17, F18, F27, F31, F36 და F39. ამრიგად, 45 პრედიქტორისაგან შეირჩა მხოლოდ 7.

2. თანმიმდებრული გამორიცხვის მეთოდი.

Feature (Independent Variable) Selection ფანჯარაში ჩავრთოდ **Backward selection** (თანმიმდებრული გამორიცხვით) ოპცია.



გავსხნათ **Interactive** (ინტერაქტიური) ჩანართი და ამოვირჩიოთ **All combinations tested** (ტესტირების ყველა კომბინაცია).



შემდეგ **Click OK** ღილაკზე.

დასაწყისში მოდელში ჩართულია ყველა პრედიქტორი და შემდეგ ყოველ ბიჯზე ხდება უმნიშვნელო პრედიქტორის გამორიცხვა. შედეგები წარმოდგენილია შემდეგ ცხრილში.

Backwards Selection (pd 2)																								
	Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23
3.2	0,066359	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.3	0,066341	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.4	0,066199	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.5	0,066549	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.6	0,065542	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.7	0,066279	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.8	0,066369	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.9	0,066349	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.10	0,066124	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.11	0,066342	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.12	0,066373	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.13	0,066374	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.14	0,066359	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.15	0,066331	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y
3.16	0,066320	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y
3.17	0,066366	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
3.18	0,066346	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y

საბოლოო შედეგი ჩაწერილია ცხრილის ბოლო სტრიქონში (*Final*).

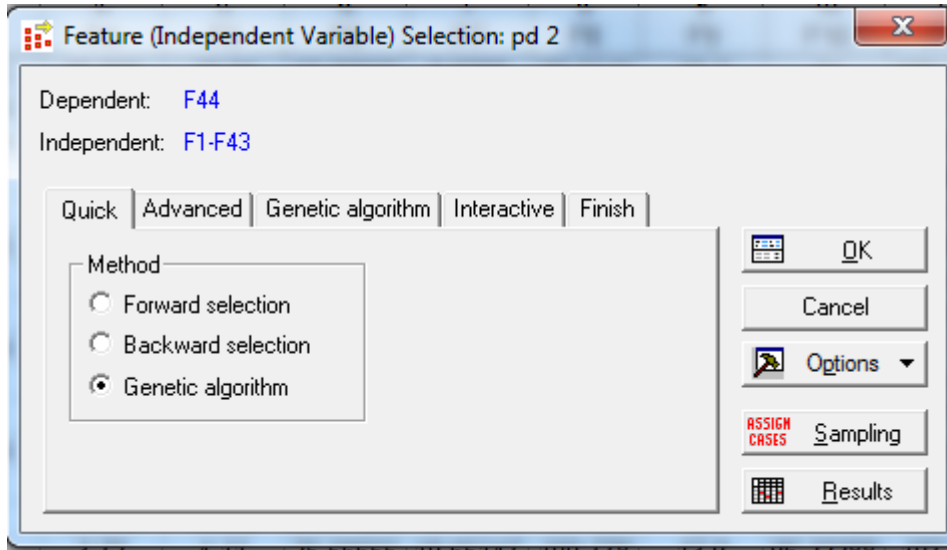
Backwards Selection (pd 2)																																												
	Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	F32	F33	F34	F35	F36	F37	F38	F39	F40			
38.33	0,062341	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	Y	Y	-	Y	Y	-	-	-	-	Y	-	
38.34	0,062403	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	Y	Y	Y	-	Y	Y	-	-	-	-	Y	-
38.35	0,062518	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	Y	Y	Y	Y	-	-	-	-	-	-	Y	-
38.39	0,062540	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	Y	Y	Y	Y	-	-	-	-	-	-	Y	-
Final	0,062324	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	Y	Y	Y	Y	-	-	-	-	-	-	Y	-

როგორც მიღებული შედეგიდან ჩანს მოდელის მიერ შერჩეული მნიშვნელოვანი ცვლადები შემდეგია: F5, F27, F31- F35 და F39.

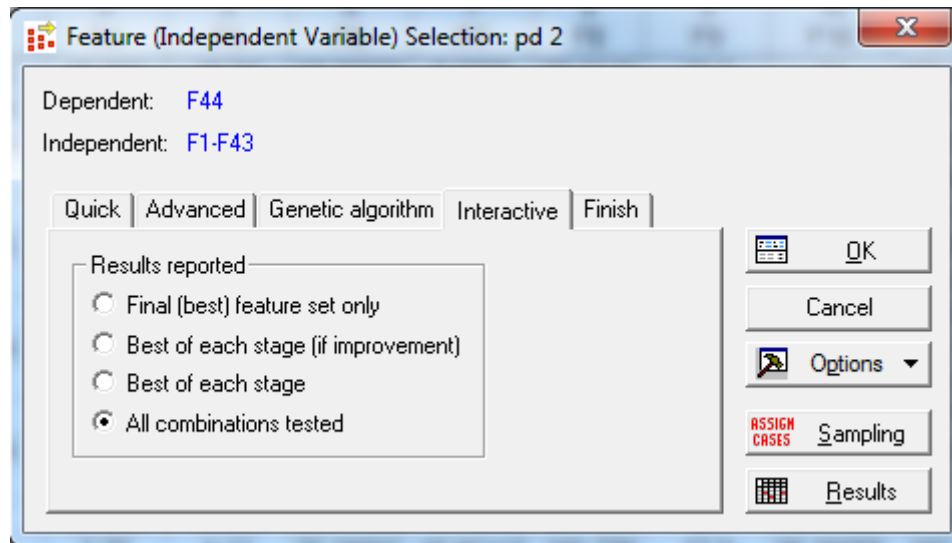
3. გენეტიკური ალგორითმი.

გენეტიკური ალგორითმი წარმოადგენს პარამეტრების შერჩევის მეტად ეფექტურ ინსტრუმენტს. მართალია ეს ალგორითმი მოითხოვს გამოთვლების დიდ დროს, მაგრამ *STATISTICA Neural Networks* პაკეტში რეალიზირებული *PNN/GRNN* ტიპის ნეიროქსელები საშუალებას იძლევიან გენეტიკური ალგორითმის მუშაობის სისწრაფვის გაზრდას.

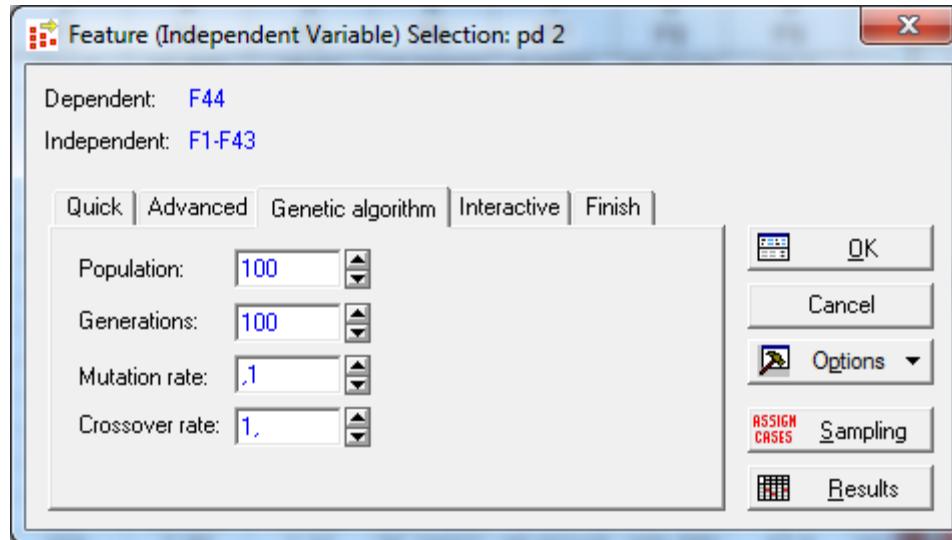
Feature (Independent Variable) Selection ფანჯარაში ჩავრთოდ *Genetic algorithm* (გენეტიკური ალგორითმი)-ის ოპცია.



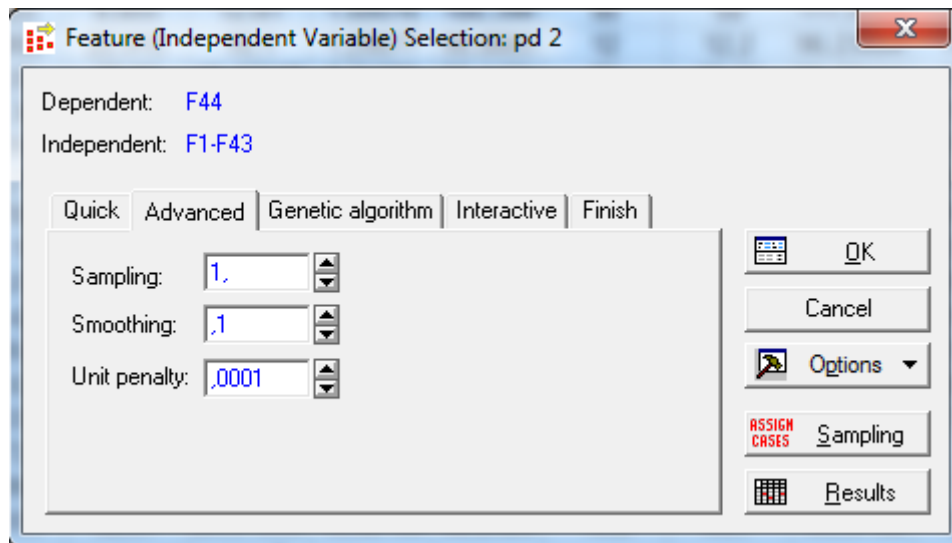
გავსნათ **Interactive** (ინტერაქტიური) ჩანართი და ამოვირჩიოთ **All combinations tested** (ტესტირების ყველა კომბინაცია).



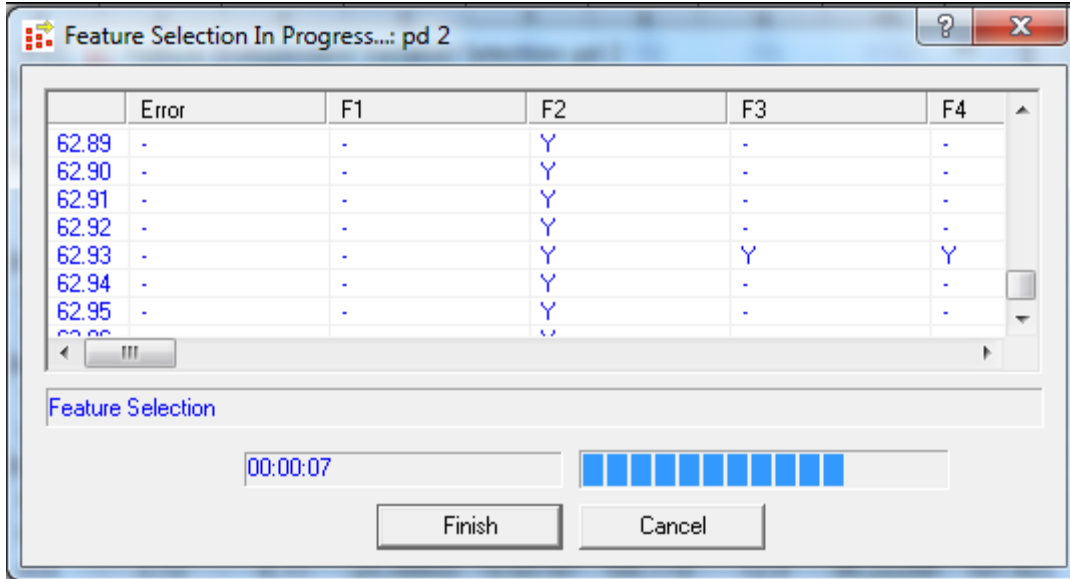
Genetic algorithm ჩანართში გაჩუმების პრინციპით წარმოდგენილია **Population** (პოპულაცია), **Generations** (მემკვიდრეობა), **Mutation rate** (მუტაციის სისწრაფე), **Crossover rate** (შეჯვარების სისწრაფე) მაჩვენებლები. დავტოვოთ ეს მაჩვენებლები.



თუ გავხსნით **Advanced** ჩანართს ვნახავთ, რომ **Unit penalty** (ჯარიმების ბლოკი)-ის ველში გაჩუმების პრინციპით ჩაწერილია 0,0001 სიდიდე. ზოგადად, რეკომენდირებულია ჯარიმის სიდიდე შევარჩიოთ 0,0001 – 0,0005 დიაპაზონში.



დავტოვოთ ეს მაჩვენებლები და შემდეგ **Click OK** დილაკზე. ეკრანზე გამოჩნდება ნეიროქსელის სწავლების დიალოგიური **Feature Selection In Progress** (მიმდინარეობს შესწავლა) ფანჯარა.



სწავლების დასრულების შემდეგ ეკრანზე გამოდის შედეგების ცხრილი

Genetic Algorithm Selection (pd 2)																								
	Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	
1.1	0,038219	Y	Y	-	-	-	Y	-	-	Y	-	-	-	Y	Y	Y	Y	-	-	-	Y	Y	Y	
1.2	0,038122	-	-	Y	Y	-	Y	Y	Y	Y	-	-	-	Y	Y	-	-	Y	-	-	Y	Y	-	Y
1.3	0,038173	Y	Y	Y	Y	Y	-	Y	Y	-	Y	-	Y	Y	Y	-	Y	-	Y	Y	Y	Y	Y	-
1.4	0,038055	-	-	Y	Y	Y	-	-	Y	Y	-	Y	-	-	-	Y	-	Y	Y	-	-	Y	-	-
1.5	0,038023	-	-	Y	-	-	Y	Y	Y	Y	-	-	-	Y	Y	-	-	Y	-	Y	Y	Y	Y	-
1.6	0,038713	Y	Y	Y	Y	Y	-	Y	Y	Y	-	-	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y
1.7	0,038036	Y	Y	-	Y	-	-	-	-	-	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-	-	-
1.8	0,037668	Y	Y	-	Y	-	-	Y	-	-	Y	-	-	-	Y	-	-	Y	-	Y	Y	Y	Y	-
1.9	0,038326	-	-	Y	Y	-	Y	Y	Y	Y	-	-	-	Y	Y	Y	-	Y	Y	-	-	Y	Y	Y
1.10	0,038472	-	-	Y	Y	-	Y	Y	Y	Y	-	-	-	Y	Y	-	-	Y	-	Y	-	-	-	-
1.11	0,037502	-	Y	-	-	-	-	-	Y	-	-	-	Y	-	-	Y	-	Y	Y	-	-	Y	Y	-
1.12	0,037960	-	-	-	Y	-	-	-	-	-	-	Y	-	Y	-	Y	-	-	Y	-	Y	-	-	-
1.13	0,037326	Y	-	-	-	-	-	Y	-	-	-	-	Y	-	-	Y	-	-	-	Y	Y	-	-	-
1.14	0,038099	-	-	Y	Y	Y	Y	Y	Y	-	-	-	-	Y	-	Y	-	-	-	Y	-	Y	-	-
1.15	0,037596	Y	-	Y	-	Y	Y	Y	-	Y	Y	-	-	Y	Y	-	-	-	Y	-	-	-	-	-
1.16	0,038646	-	Y	Y	Y	Y	Y	Y	-	Y	-	Y	Y	Y	-	-	Y	-	-	Y	-	Y	-	-

საბოლოო შედეგი ჩაწერილია ცხრილის ბოლო სტრიქონში (*Final*).

Genetic Algorithm Selection (pd 2)																																														
	Error	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	F32	F33	F34	F35	F36	F37	F38	F39	F40	F41	F42			
101.41	0,035973	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	Y	-	-	Y	-	-	-	Y	-	Y	-	-		
101.42	0,036011	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	Y	-	-	Y	-	-	-	Y	-	Y	-	-	-	
101.43	0,036020	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	Y	-	-	Y	-	-	-	Y	-	Y	-	-	-	-
Final	0,035929	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	-	-	Y	-	-	Y	-	-	-	Y	-	Y	-	-	-	-

როგორც მიღებული შედეგიდან ჩანს მოდელის მიერ შერჩეული მნიშვნელოვანი ცვლადები შემდეგია: F3, F6, F10, F26, F27, F31, F35 და F39.

ლიტერატურა

1. Нейронные сети. *STATISTICA Neural Networks*. Методология и технология современного анализа данных. Под редакцией В.П.Боровикова. М.,Горячая линия – Телеком, 2008.
2. Жильцов В.В.б Чувикова В.В. Практикум по нейросетевым технологиям. Учебно – методическое пособие. Омск, СибАДИ, 2010.
3. *Ясницкий Л.Н.* Введение в искусственный интеллект: учебное пособие для студ. высш. учеб. заведений. – М.: Издательский центр «Академия», 2005
4. *Башмаков А.И., Башмаков И.А.* Интеллектуальные информационные технологии: учебное пособие.– М.: Изд-во МГТУ им.Баумана, 2005.
5. ე. ყუბანეიშვილი ხელოვნური ნეირონული ქსელები მედიცინაში. ლექციების კურსი. სტუ, 2013, [ihhttp://gtu.ge/books/ ms/xel_medicina.pdf](http://gtu.ge/books/ms/xel_medicina.pdf).