ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ

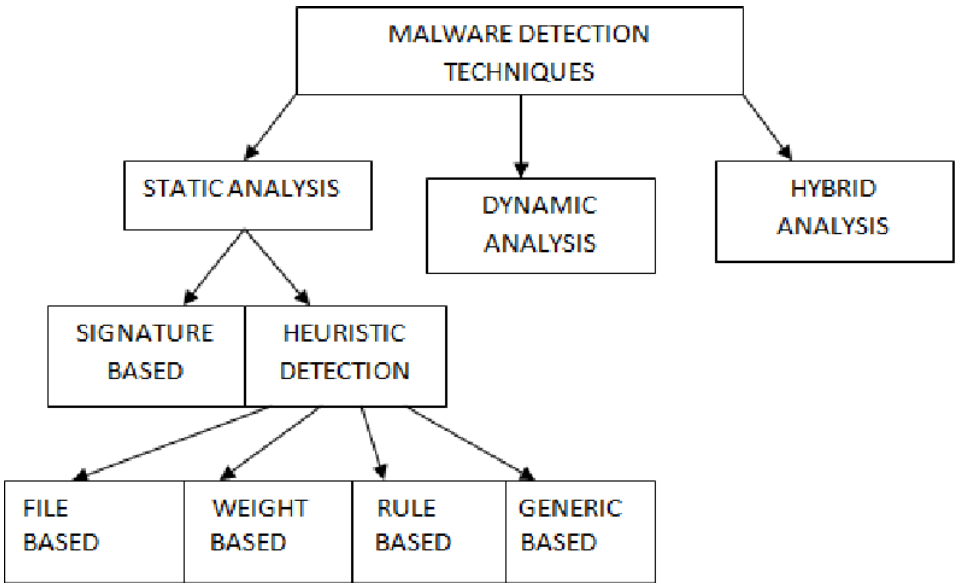
M.Sc IN APPLIED INFORMATICS

XLCNN: PRE-TRAINED TRANSFORMER MODEL FOR MALWARE DETECTION

Konstantinos Giapantzis

ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
University of Macedonia
MSc in Applied Informatics

Malware Analysis Techniques



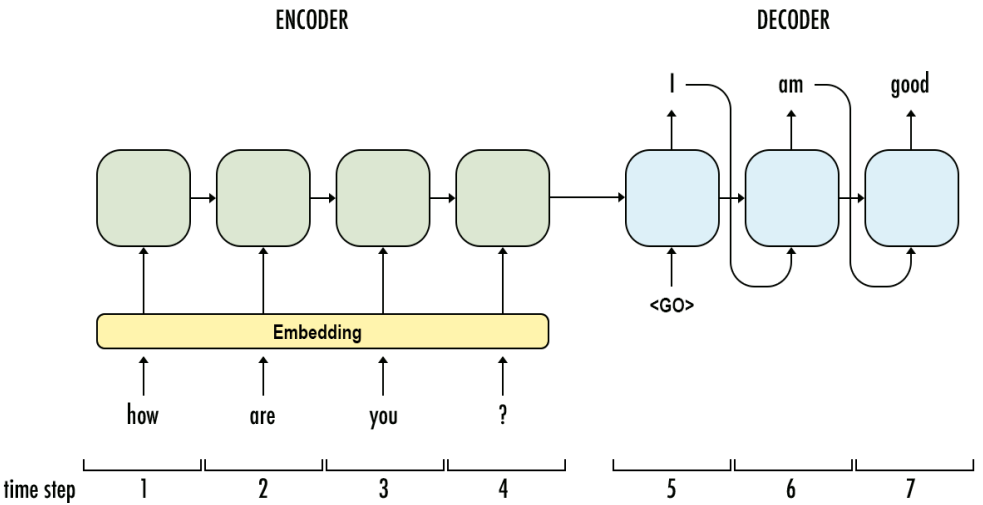
Malware Types

- Virus
- Trojans
- Worms
- Adware

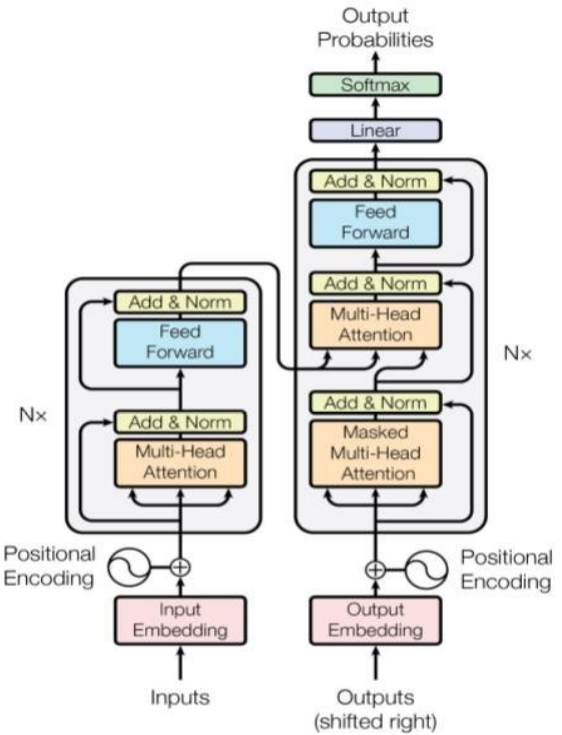
- Rootkit
- Bots
- Ransomware

Sequence-to-Sequence

- **Embedding layer:** maps the input words
- **Encoder Layer:** compresses the tensors to a predetermined size
- **Decoder layer:** produces the output token sequence



Attention Mechanism



XLCNN Architecture

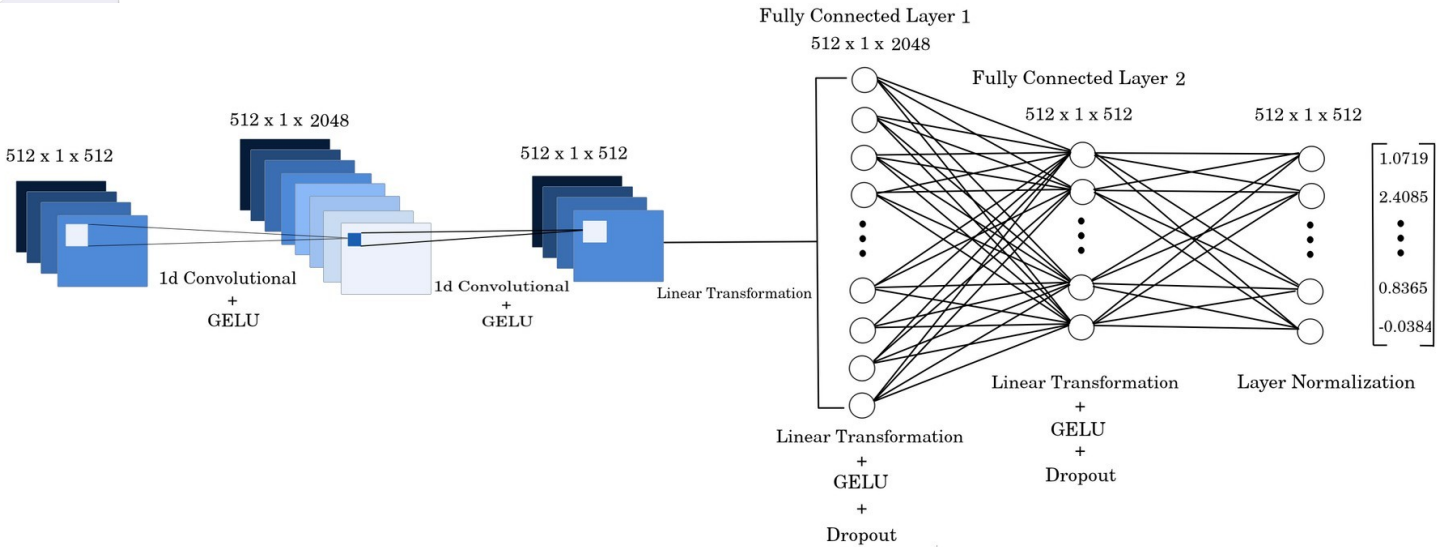
XLCNN uses embeddings with $32 * 10^3$ token vocabulary

- CNN Layer 1
  Input matrix: [512 x 1 x 512]
  Output matrix: [512 x 1 x 2048]

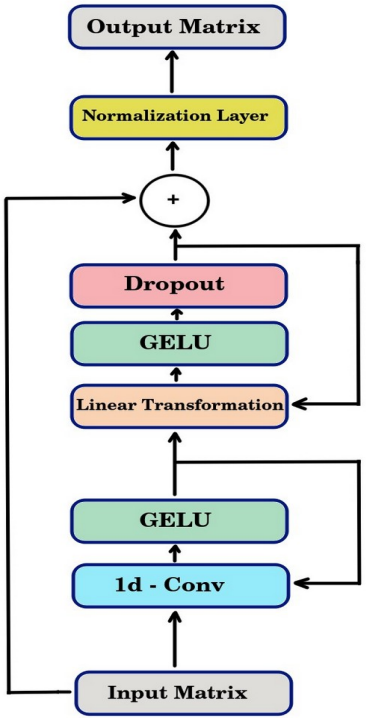
- CNN Layer 2
  Input matrix: [512 x 1 x 2048]
  Output matrix: [512 x 1 x 512]

- Feed Forward Layers
  24 layers
  Input matrix: [512 x 1 x 512]
  Output matrix: [512 x 1 x 512]

XLCNN Architecture



XLCNN Feed Forward Network



Normalization: $$y = \frac{x - E[x]}{\sqrt{Var[x] + \varepsilon}} \cdot \gamma + \beta$$

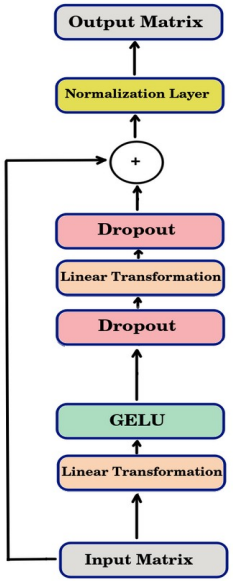
Dropout: f(x, p) = $$\begin{cases} x, if\ x \geq p \\ 0, if\ x < p \end{cases}$$

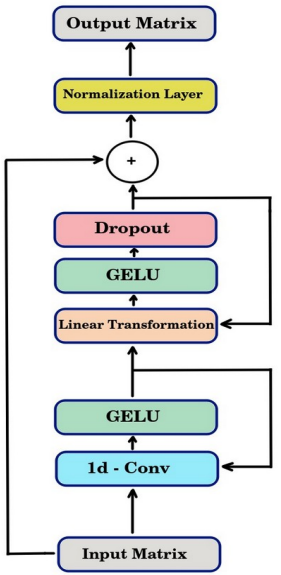
GELU: $$G(x) = MAD(x) \cdot \int_0^x e^{\frac{-(x-\mu)^2}{2 \cdot \sigma^2}} dx$$

Linear Trans: $$y = G(x) \cdot w^T + b$$

CNN: $$C(x) = \beta \cdot b_m + a \cdot (x \cdot w_m)$$

XLNet vs XLCNN FF Architecture





Data Source Collection – Benign data

Windows 7: 342

Windows 8: 608

Windows 10: 950

Total Benign data: 1794

Training data: 1641

Testing data: 153

Data Source Collection – Malicious data

- **VirusTotal:** online scan engine to check for viruses, owned by Chronicle
- **Das Malwerk:** online repository owned by Robert Svensson, security researcher at FlixBus company

Total malicious data: 3117

Training malicious data: 2824

Testing malicious data: 293

Metadata extraction

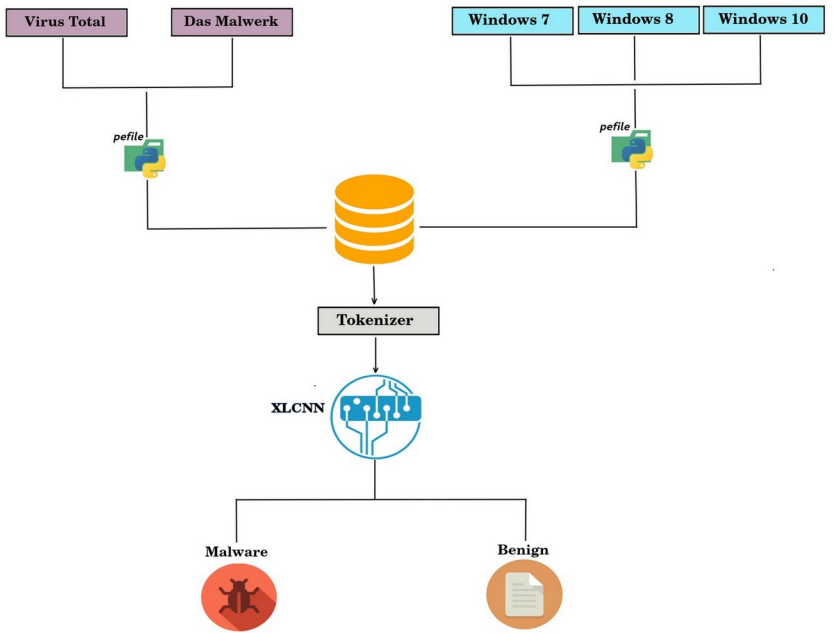
- The metadata in this research project was extracted from the executable Windows files using a python library called pefile
- Pefile is a multi-platform Python module to parse and work with Portable Executable (PE) files
- Most of the information contained in the PE file headers is accessible as well as all the sections details and data

Metadata from Windows executables files

```
[IMAGE_IMPORT_DESCRIPTOR]
0x46E8     0x0   OriginalFirstThunk:           0x4724
0x46E8     0x0   Characteristics:              0x4724
0x46EC     0x4   TimeDateStamp:                0x0        [Thu Jan  1 00:00:00 1970 UTC]
0x46F0     0x8   ForwarderChain:               0x0
0x46F4     0xC   Name:                         0x48A2
0x46F8     0x10  FirstThunk:                   0x2000

AVIFIL32.dll.AVIStreamCreate Hint[36]

[IMAGE_IMPORT_DESCRIPTOR]
0x46FC     0x0   OriginalFirstThunk:           0x4770
0x46FC     0x0   Characteristics:              0x4770
0x4700     0x4   TimeDateStamp:                0x0        [Thu Jan  1 00:00:00 1970 UTC]
0x4704     0x8   ForwarderChain:               0x0
0x4708     0xC   Name:                         0x48BE
0x470C     0x10  FirstThunk:                   0x204C

WinSCard.dll.SCardCancel Hint[6]

---------Resource directory---------

[IMAGE_RESOURCE_DIRECTORY]
0x14000    0x0   Characteristics:              0x0
0x14004    0x4   TimeDateStamp:                0x0        [Thu Jan  1 00:00:00 1970 UTC]
0x14008    0x8   MajorVersion:                 0x0
0x1400A    0xA   MinorVersion:                 0x0
0x1400C    0xC   NumberOfNamedEntries:         0x0
0x1400E    0xE   NumberOfIdEntries:           0x3
  Id: [0x3] (RT_ICON)
  [IMAGE_RESOURCE_DIRECTORY_ENTRY]
  0x14010    0x0   Name:                       0x3
  0x14014    0x4   OffsetToData:               0x80000028
    [IMAGE_RESOURCE_DIRECTORY]
    0x14028    0x0   Characteristics:          0x0
    0x1402C    0x4   TimeDateStamp:            0x0        [Thu Jan  1 00:00:00 1970 UTC]
    0x14030    0x8   MajorVersion:             0x0
    0x14032    0xA   MinorVersion:             0x0
    0x14034    0xC   NumberOfNamedEntries:     0x0
    0x14036    0xE   NumberOfIdEntries:       0x6
      Id: [0x1]
      [IMAGE_RESOURCE_DIRECTORY_ENTRY]
      0x14038    0x0   Name:                   0x1
      0x1403C    0x4   OffsetToData:           0x800000A8
        [IMAGE_RESOURCE_DIRECTORY]
        0x140A8    0x0   Characteristics:      0x0
        0x140AC    0x4   TimeDateStamp:        0x0        [Thu Jan  1 00:00:00 1970 UTC]
        0x140B0    0x8   MajorVersion:         0x0
        0x140B2    0xA   MinorVersion:         0x0
        0x140B4    0xC   NumberOfNamedEntries: 0x0
        0x140B6    0xE   NumberOfIdEntries:    0x1
        \--- LANG [9,1][LANG_ENGLISH,SUBLANG_ENGLISH_US]
          [IMAGE_RESOURCE_DIRECTORY_ENTRY]
```

```
USER32.dll.DestroyCursor Hint[162]
USER32.dll.SetWindowPos Hint[710]
USER32.dll.DestroyWindow Hint[166]
USER32.dll.LoadStringA Hint[505]
USER32.dll.LoadIconA Hint[492]
USER32.dll.IsWindowEnabled Hint[476]
USER32.dll.CharLowerW Hint[46]
USER32.dll.GetForegroundWindow Hint[301]
USER32.dll.GetKeyboardLayout Hint[318]
USER32.dll.IsWindowVisible Hint[480]
USER32.dll.EnableWindow Hint[216]
USER32.dll.ScreenToClient Hint[621]
USER32.dll.LoadCursorW Hint[491]
USER32.dll.DefWindowProcW Hint[156]
USER32.dll.TrackMouseEvent Hint[757]
USER32.dll.ShowWindow Hint[735]
USER32.dll.GetClassNameW Hint[274]
USER32.dll.MapWindowPoints Hint[521]
USER32.dll.CharNextW Hint[49]
USER32.dll.GetParent Hint[356]
USER32.dll.PtInRect Hint[576]
USER32.dll.TranslateAcceleratorW Hint[762]
USER32.dll.MessageBeep Hint[525]
USER32.dll.UpdateLayeredWindow Hint[782]
USER32.dll.LoadStringW Hint[506]

[IMAGE_IMPORT_DESCRIPTOR]
0x1441C    0x0   OriginalFirstThunk:           0x16B84
0x1441C    0x0   Characteristics:              0x16B84
0x14420    0x4   TimeDateStamp:                0x0        [Thu Jan  1 00:00:00 1970 UTC]
0x14424    0x8   ForwarderChain:               0x0
0x14428    0xC   Name:                         0x18206
0x1442C    0x10  FirstThunk:                   0x1408C

GDI32.dll.CloseFigure Hint[30]
GDI32.dll.BeginPath Hint[18]
GDI32.dll.AddFontMemResourceEx Hint[2]

[IMAGE_IMPORT_DESCRIPTOR]
0x14430    0x0   OriginalFirstThunk:           0x16AF8
0x14430    0x0   Characteristics:              0x16AF8
0x14434    0x4   TimeDateStamp:                0x0        [Thu Jan  1 00:00:00 1970 UTC]
0x14438    0x8   ForwarderChain:               0x0
0x1443C    0xC   Name:                         0x184CA
0x14440    0x10  FirstThunk:                   0x14000

ADVAPI32.dll.SetTokenInformation Hint[706]
ADVAPI32.dll.CreateServiceW Hint[129]
ADVAPI32.dll.OpenServiceW Hint[507]
ADVAPI32.dll.RegisterEventSourceW Hint[643]
ADVAPI32.dll.RegOpenKeyA Hint[607]
ADVAPI32.dll.SetNamedSecurityInfoW Hint[689]
ADVAPI32.dll.RegQueryValueExW Hint[622]
```

Application on malware detection

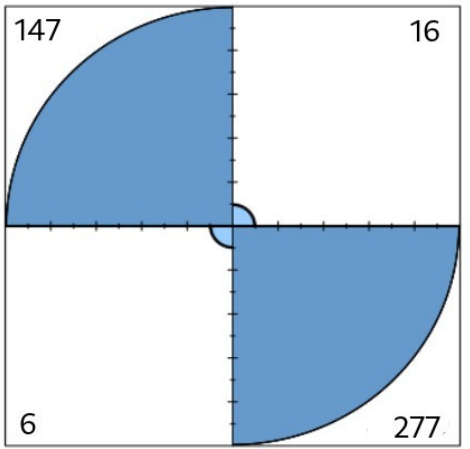


Total training data: 4452

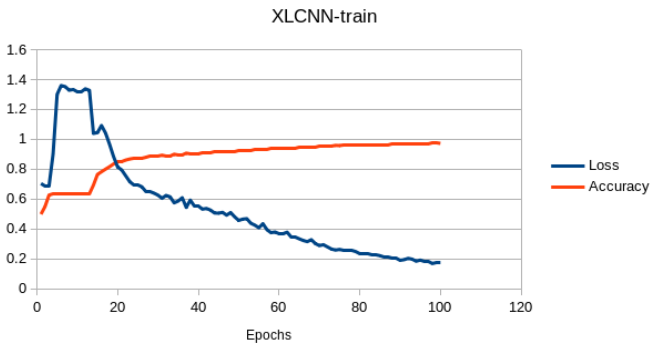
Total testing data: 446

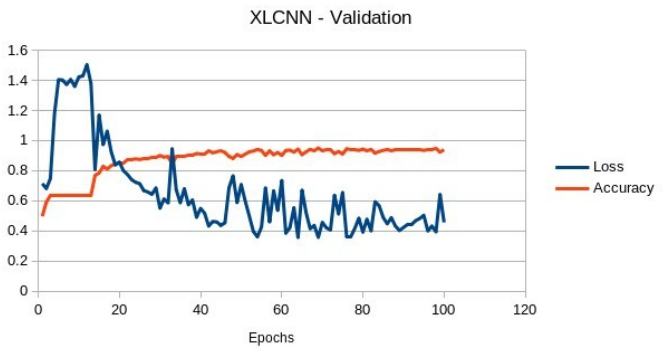
XLCNN Results - Decision

Desicion	XLCNN Transformer	
	Hypothesis True	**Hypothesis False**
Malware	277	16
Benign	147	6



XLCNN Results - Graphs

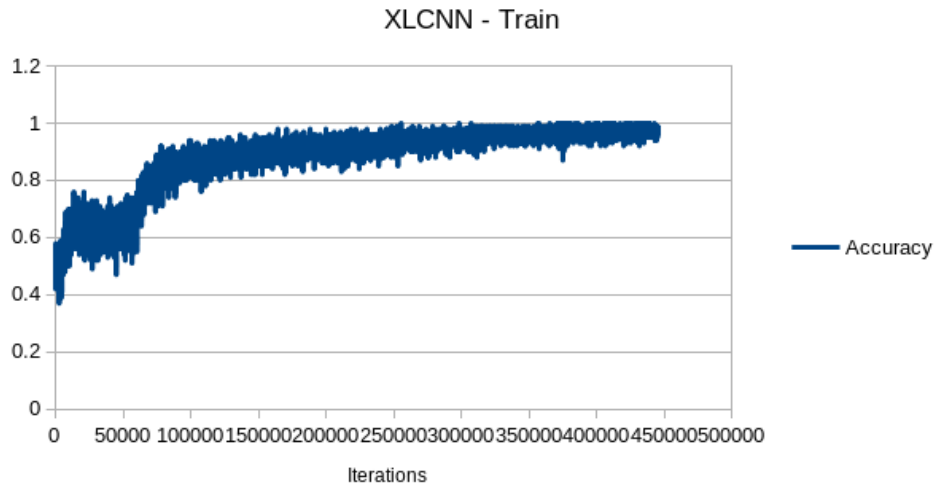




XLCNN Results - Graphs

Epochs: 100
Iteration per epoch: 4452
Total iterations: 445200

Trainning time: 8.3 epochs/hour
Total training time: 120 hours



Comparison with XLNet

	recall	FPR	TNR	FNR	precision	NPV	FDR	f1	AUC	accuracy
XLCNN	97.88	9.82	90.18	2.12	94.54	96.08	5.46	96.18	94.03	95.07
XLNet	96.82	12.27	87.73	3.18	93.19	94.08	6.80	94.97	92.27	93.49

Limitations

- Small amount of training data. The model according to the results obtained could achieve an even greater success rate in its prediction, if it had more data for its training.
- Input tensors size. This limitation has to do with the memory of the graphics card used to calculate the mathematical operations.
- Feed forward network size

Future Extensions

The deep learning technique used for XLCNN belongs to the category of supervised learning.

A future expansion could be the creation of a network to detect malware without
supervision.

A deep reinforcement learning algorithm like deep q learning or actor critic reinforcement learning could be applied, which would take the XLCNN as a pre-trained model.

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ

Thank you for your attention!

Konstantinos Giapantzis

ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
University of Macedonia
MSc in Applied Informatics