



**ΜΑΘΗΤΙΚΟ ΣΥΝΕΔΡΙΟ ΠΛΗΡΟΦΟΡΙΚΗΣ  
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ**

2009-2018 Δέκα Χρόνια Μαθητικής Ψηφιακής Καινοτομίας

# Σύνθεση μουσικής με τεχνητή νοημοσύνη

Παναγιώτης Πετρίδης & Οδυσσέας Σαποζνίκοβ  
1ο Γενικό Λύκειο Αλεξάνδρειας

# Πως σκεφτήκαμε αυτή την ιδέα;

Η τεχνητή νοημοσύνη είναι στο αποκορύφωμα της στις μέρες μας. Αυτοκίνητα που οδηγούν μόνα τους, προγράμματα που κάνουν διάγνωση σε ασθενείς, προτάσεις προϊόντων, η τεχνητή νοημοσύνη είναι παντού.

Έχοντας δουλέψει ξανά με τα συστήματα νευρώνων στην περσινή μας εργασία και γνωρίζοντας τι μπορούν να κάνουν σκεφτήκαμε ότι θα μπορούσαν να γράφουν και μουσική.





# Δομή της παρουσίασης

1. Συλλογή δεδομένων
2. Επεξεργασία δεδομένων
3. Το νευρωνικό δίκτυο
4. Αποτελέσματα & Demo

# Συλλογή Δεδομένων

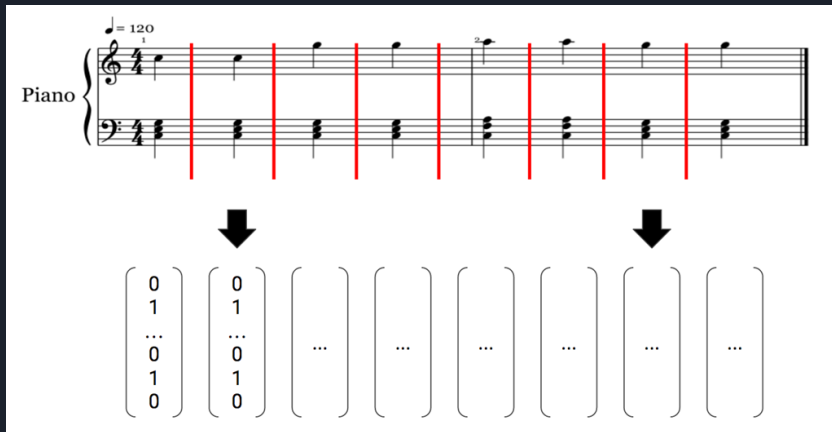
```
X: 5
T:Bobbin Mill Reel
% Nottingham Music Database
S:Mick Peat
M:4/4
L:1/4
K:D
A/2|:"D"d/2c/2d/2e/2 fA|"G"Be "A"cA|"D"d/2c/2d/2e/2 f/2e/2d/2c/2|"Em"Be "A"aA|
"D"d/2c/2d/2e/2 f/2e/2d/2c/2|"G"Be "A"cA|"G"B/2c/2d/2B/2 "A"c/2d/2e/2c/2|\
"D"d2 dA:|
"G"B/2A/2B/2c/2 "D"dA|"G"B/2A/2B/2c/2 "D"dA|\
"G"B/2A/2B/2c/2 "D"d/2c/2d/2e/2|"E"f/2e/2d/2f/2 "A"ea|"G"B/2A/2B/2c/2 "D"dA|
"G"B/2A/2B/2c/2 "D"de/2f/2|"G"g/2f/2e/2g/2 "D"f/2e/2d/2f/2|1\
"E"e/2d/2c/2d/2 "A"e2:|2"A"e/2d/2c/2e/2 "D"d2|
```

Παράδειγμα αρχείου ABC

Τα συστήματα νευρώνων προκειμένου να λειτουργήσουν χρειάζονται μεγάλο όγκο από δεδομένα. Στην περίπτωση μας κομμάτια μουσικής. Σε τι μορφή όμως θα πρέπει να τα δώσουμε στο νευρωνικό δίκτυο; Τα αρχεία mp4 δεν μας λένε πολλά για τις νότες, τα αρχεία WAV μπορεί να λειτουργήσουν άλλα θα είναι δύσκολο για το νευρωνικό δίκτυο να ξεχωρίσει τις νότες βασισμένο μόνο στην συχνότητα του ήχου.

Χρειαζόμαστε ένα είδος αρχείου το οποίο θα αποτελείται από τις νότες τις μουσικής. Αυτή η αναζήτησή μας για την κατάλληλη μορφή αρχείου μας οδήγησε στα αρχεία ABC τα οποία έχουν γραμμένα τα κομμάτια μουσικής με νότες για κάθε μουσικό όργανο και χωρίς τη φωνή κάποιου τραγουδιστή.

# Επεξεργασία δεδομένων



One-hot encoding

Το πρόβλημα που έχουμε ως τώρα “ένα πρόγραμμα που να γράφει μουσική” είναι πολύ αυθαίρετο. Χρειαζόμαστε κάτι πιο συγκεκριμένο.

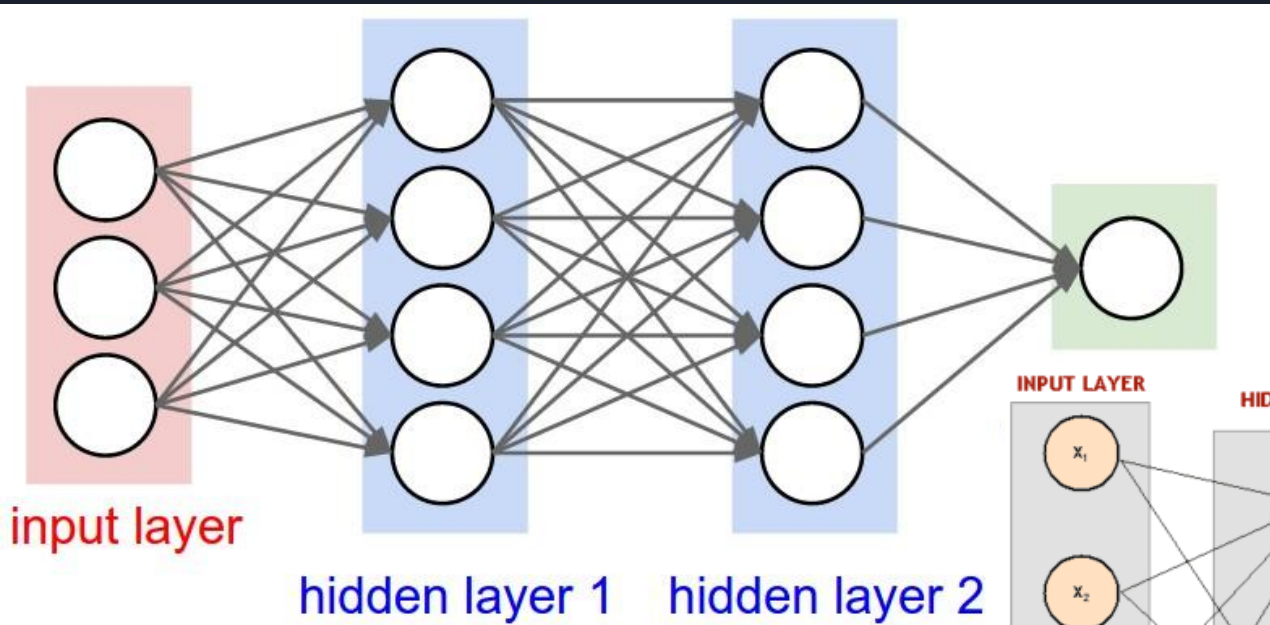
Μπορούμε να θέσουμε το πρόβλημα στο εξής: “Έχοντας μια σειρά από προηγούμενες νότες, ποια είναι η πιθανότητα ότι η νότα  $X$  θα είναι επόμενη;”.

Για παράδειγμα αν δούμε ότι μετά από το μοτίβο AGCC ακολουθεί πάντα η νότα D τότε η D έχει μεγάλη πιθανότητα να είναι επόμενη.

# ΘΕΩΡΗΤΙΚΌ ΥΠΌΒΑΘΡΟ

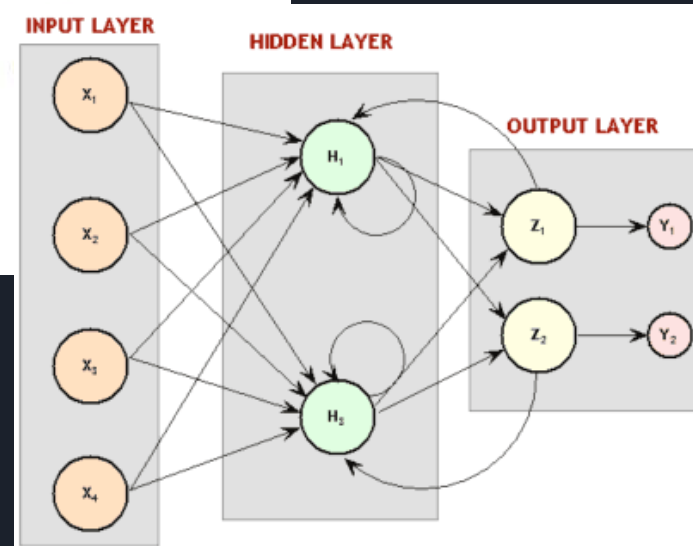
$bc \operatorname{ch} z$   $a+c=b+d$   $x \operatorname{Arth} t = \ln\left(\frac{1+t}{1-t}\right)$   $\operatorname{ch} z = \frac{e^z + e^{-z}}{2}$   
 $(ad-bc)$   $(-1 < t < 1)$   $\operatorname{sh} x = \frac{2t}{1-t^2}$   $\prod_{i=1}^n y_i$   
 $\prod_{i=1}^n y_i$   $h_1$   $h_2$   $4 \cos \omega t$ ,  $t = \frac{\pi}{\omega}$   
 $\prod_{i=1}^m y_{n+i}$   $s$   $(a-b)(c-d) = (ac+bd)$   
 $\operatorname{ch} x = \frac{1}{2} \operatorname{sh} 2x$   $\prod_{i=1}^m y_{n+i}$   $0, t < \frac{\pi}{\omega}$   
 $\frac{2dt}{1-t^2}$ ;  $\operatorname{sh}^2 x = \frac{1}{2}(\operatorname{ch} 2x - 1)$   
 $p^r = i$   $\operatorname{ch}^2 x = \frac{1}{2}(\operatorname{ch} 2x + 1)$   
 $\sum_{x=1}^m (a_x b_x)$   $\operatorname{th} \frac{\pi}{2} = t$ ,  $-4$   
 $x) = \frac{f(\mu)}{f(\beta)} = \frac{s}{k}$ ;  $\sum_{i=1}^d x_i + \sum_{i=1}^n x_{n+i}$   $\int f(x,y,z) dz$   $x \operatorname{Arth} t =$   
 $\operatorname{ch}^2 x \cdot \operatorname{sh}^2 x = 1$ ;  $a+c$

# Με μια ματιά



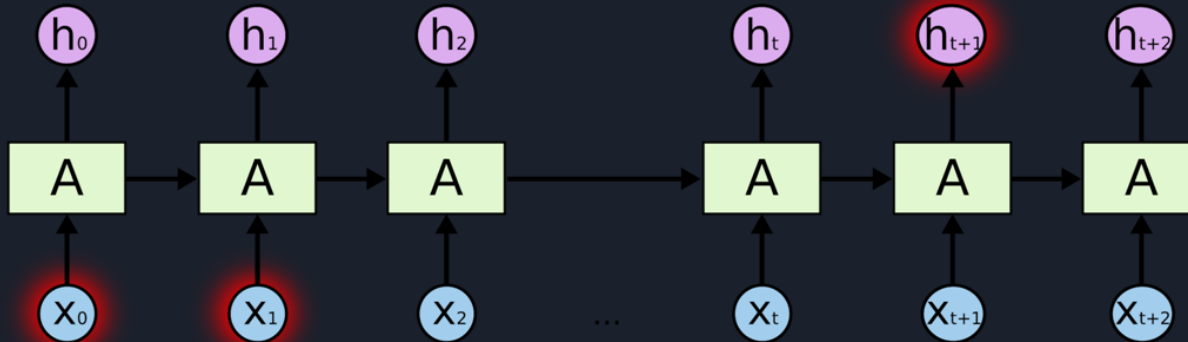
Feed Forward Neural Network

Recurrent Neural Network



# Το νευρωνικό δίκτυο

Ένα απλό Feed Forward Neural Network θα έβλεπε άπλα την αμέσως προηγούμενη νότα και θα προέβλεπε την επόμενη. Αντίθετα χρειαζόμαστε μια αρχιτεκτονική δικτύου που θα μας επιτρέψει να “θυμόμαστε” τι νότες είχαμε δει πιο πριν. Την λύση σε αυτό ακριβώς το πρόβλημα έχουν δώσει τα LSTM ( Long Short Term Memory ) Networks, επίσης γνωστά ως δίκτυα μεγάλης βραχυπρόθεσμης μνήμης. Ουσιαστικά όμως μπορούν να “μεταφέρουν” τις προβλέψεις στην επόμενη σειρά δεδομένων για να βοηθήσουν το δίκτυο να αποκτήσει έξτρα πληροφορίες.





# Αποτελέσματα & Demo



“Εκπαιδεύοντας” το δίκτυό μας για αρκετές ώρες μπορούμε πλέον να το χρησιμοποιήσουμε για να γράφουμε μουσική. Για να το κάνουμε αυτό θα ξεκινήσουμε από μία τυχαία σειρά από νότες και θα αφήσουμε το πρόγραμμα να συνεχίσει το μοτίβο με νότες έτσι ώστε να ακούγεται όμορφα. Πως ακούγεται; Μπορείτε να κρίνετε και μόνοι σας! Αποτελέσματα εδώ:

[https://soundcloud.com/panos\\_ptr/sets/deepnote-predictions](https://soundcloud.com/panos_ptr/sets/deepnote-predictions)



# Αναφορές

1. Oliwa, T., & Wagner, M. (2008). Composing music with neural networks and probabilistic finite-state machines. [10.1007/978-3-540-78761-7\\_55](https://doi.org/10.1007/978-3-540-78761-7_55)
2. Sak, H., Senior, A., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. Paper presented at the Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 338-342.
3. Guerzhoy M. (2016). One hot encoding. Slides on “Intro to Machine Learning and Neural Networks” available online:  
<http://www.cs.toronto.edu/~guerzhoy/321/lec/W04/onehot.pdf>

Σας ευχαριστούμε!