

Searchable handwritten notes using Edge-AI



2024-09-26

MLOps Scotland

Martin Lellep

Get a copy of these slides here



<https://lellep.xyz/blog/slides-for-my-mlops-2024-talk>

What I want to you take from here

Understand ...

... an example for
a FOSS ML
feature
contribution

MLOps ...

... considerations
for Edge-AI FOSS
use cases

Actionable ...

... tips for you
when looking for a
project to
contribute to

FOSS = Free and Open-Source Software

What is Edge-AI?

Edge artificial intelligence refers to the deployment of AI algorithms and AI models directly on local edge devices such as sensors or **Internet of Things** (IoT) devices, which enables real-time data processing and analysis without constant reliance on cloud infrastructure.

- Definition of IBM

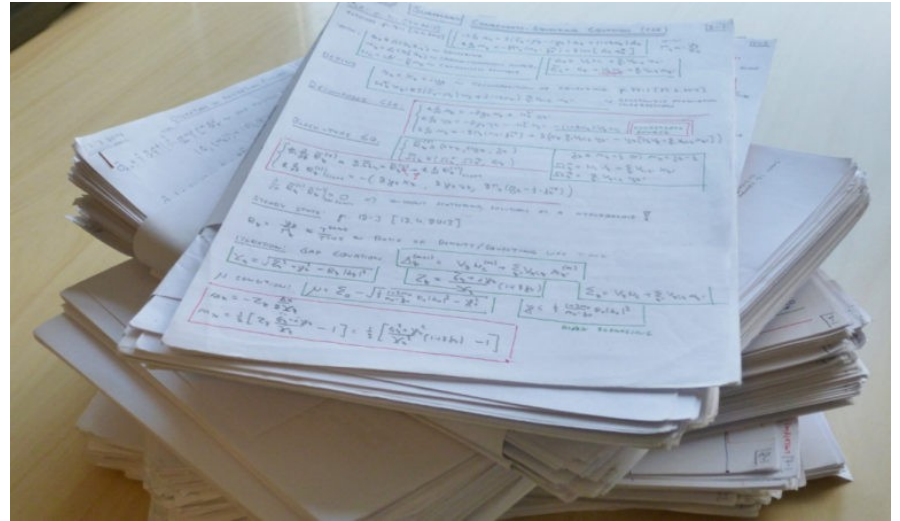
I am Martin Lellep



Experience: Applied Science, Machine Learning

- Just finished PhD: computational + theoretical physics with ML
- Previously:
 - Finance research
 - GenAI for healthcare
at Canon Medical Research Europe in Edinburgh

I take a lot of handwritten notes!



Study notes & regular notes & maths notes

Digital handwritten note-taking

People like to take digital handwritten notes

Benefits:

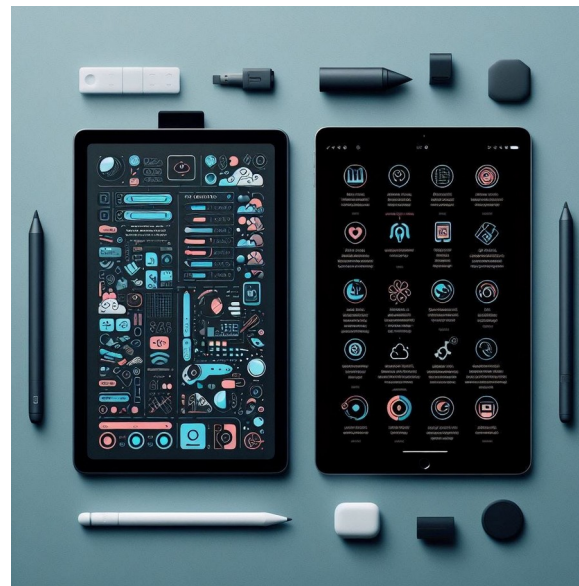
- Reduced paper consumption
- Fully editable
- Fully searchable

Fosters digitalisation and thereby businesses (e.g. RAG)

Example: GoodNotes on iPad

- 24 million MAU (source)
- ARR: O(millions) (source)

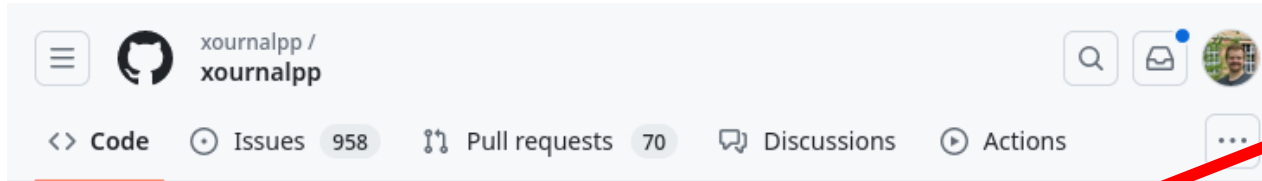
**Problem:
vendor lock-in**



source

Xournal++

FOSS handwritten notes cross-platform application



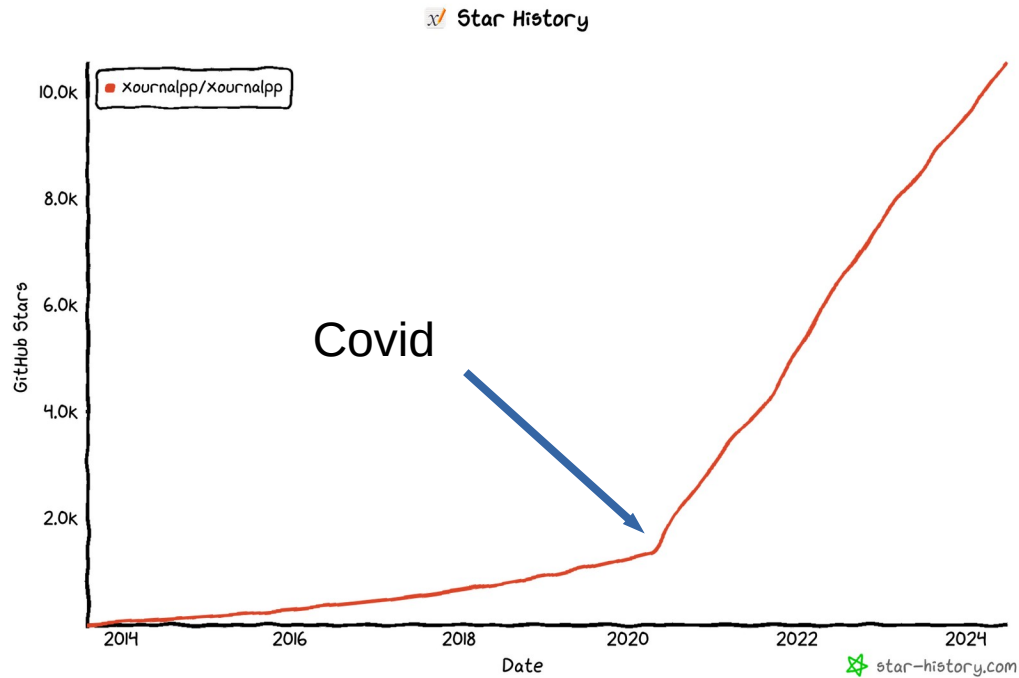
10.5k stars

Open file format

*Your notes
truly belong
you!*

source

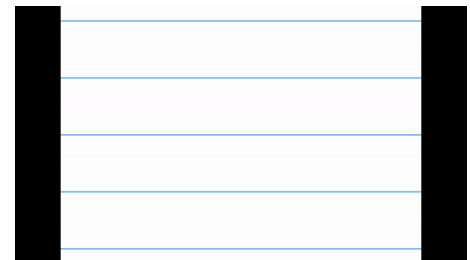
Xournal++ grows!



Xournal++ is written in C++

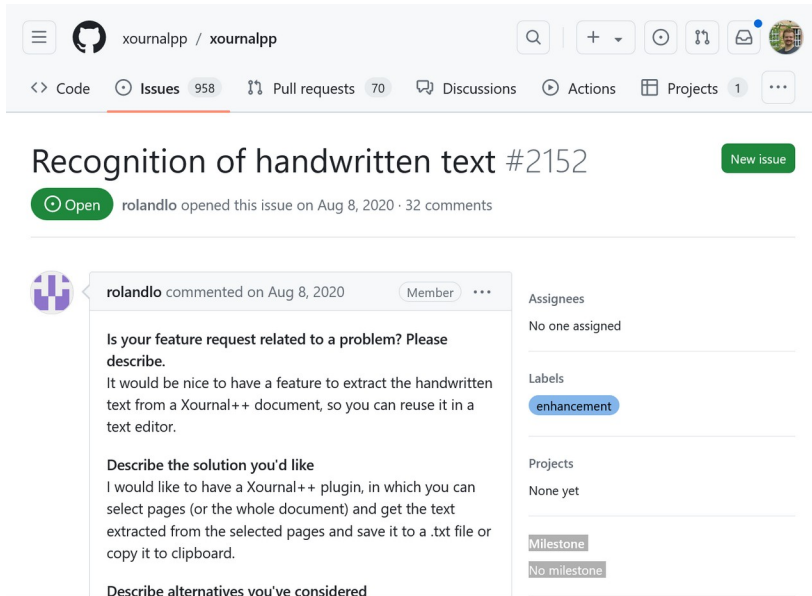
My first production contributions:

1. Doxygen for docs
2. Coordinate system tool



Problem: Not searchable

We cannot search our Xournal++ notes!



The screenshot shows a GitHub issue page for the repository 'xournalpp / xournalpp'. The issue title is 'Recognition of handwritten text #2152', marked as a 'New issue'. It was opened by 'rolandlo' on August 8, 2020, and has 32 comments. A comment from 'rolandlo' is visible, asking for a feature to extract handwritten text from Xournal++ documents. The comment text is: 'Is your feature request related to a problem? Please describe. It would be nice to have a feature to extract the handwritten text from a Xournal++ document, so you can reuse it in a text editor. Describe the solution you'd like I would like to have a Xournal++ plugin, in which you can select pages (or the whole document) and get the text extracted from the selected pages and save it to a .txt file or copy it to clipboard. Describe alternatives you've considered'. The issue is labeled as an 'enhancement' and has no assignees or projects.

Community wants this feature!

But not via cloud.

Challenges for search implementation:

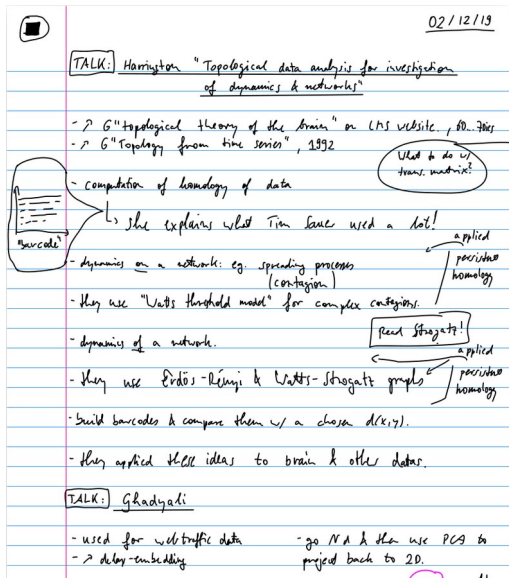
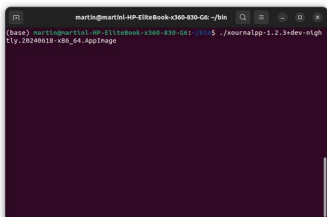
- Local to be privacy aware
- Reasonably fast
- Reasonably accurate
- Run on limited hardware
- Callable from C++
- Delivered to edge
- Make community aware of solution
- Community not skilled in ML

 = MLOps questions!

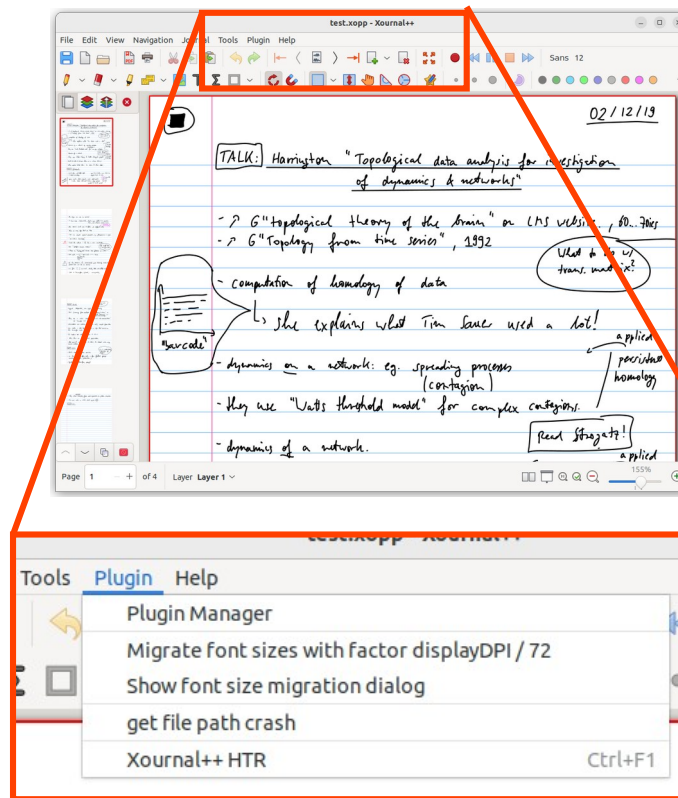
My solution: Xournal++ HTR

HTR = Handwritten Text Recognition

Start
(currently
nightly)



Result:
PDF



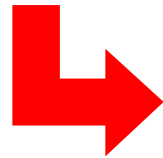
Call plugin

Select
destination

My solution: Xournal++ HTR

HTR = Handwritten Text Recognition

Resulting
searchable
PDF



Searchable
through
(invisible)
overlaid
text

02/12/19

TALK: Hamington Topological data analysis for investigation of dynamics & networks

-> G "topological theory of the brain" on LHS website, 00... Jries
-> G "topology from time series, 1992

What to do w/ trans. matrix?

computation of homology of data
She explains what Tim Lane used a dot!

dynamics on a network: eg spreading process (contagion)
- they use "Virus threshold model" for complex contagions.

Real Strazats!

- they use Erdos-Renyi & Watts-Strazats graphs
applied persistence homology

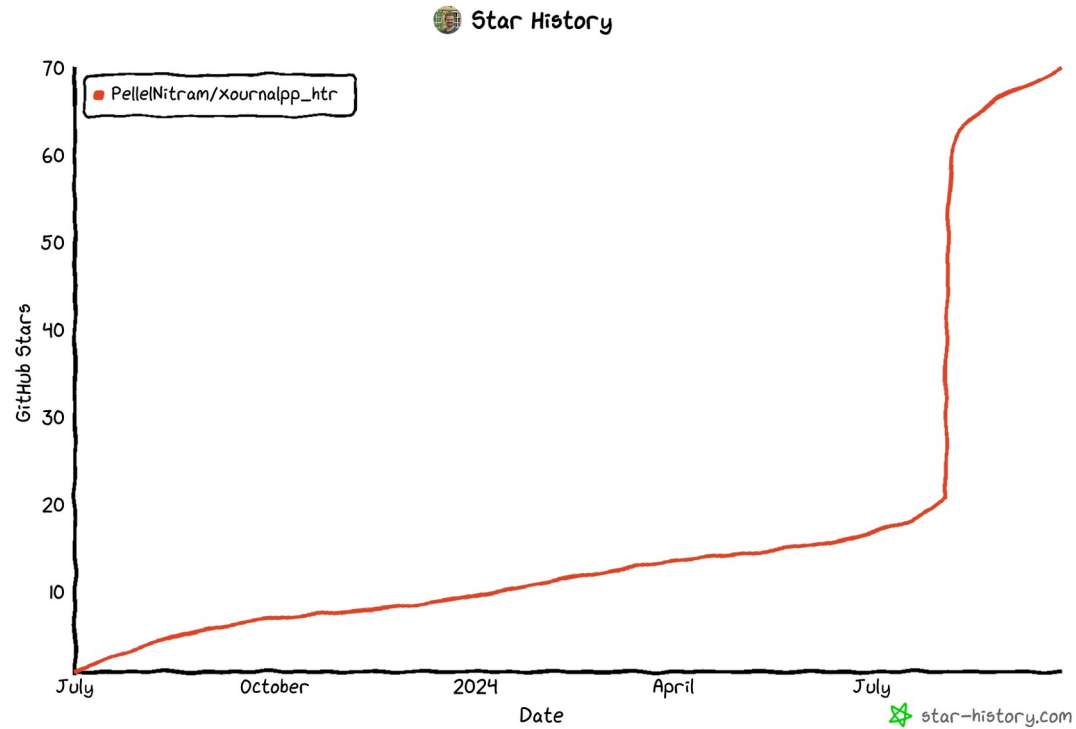
build barcodes & compare them w/ a chosen $d(x,y)$.

- they applied VLL ideas to brain & other data.

TALK: Ghadyani

- used for web traffic data
-> delay-embedding
- go NA & then use PCA to project back to 2D.

Xournal++ HTR grows



Community seems to like it!

Live Demo!

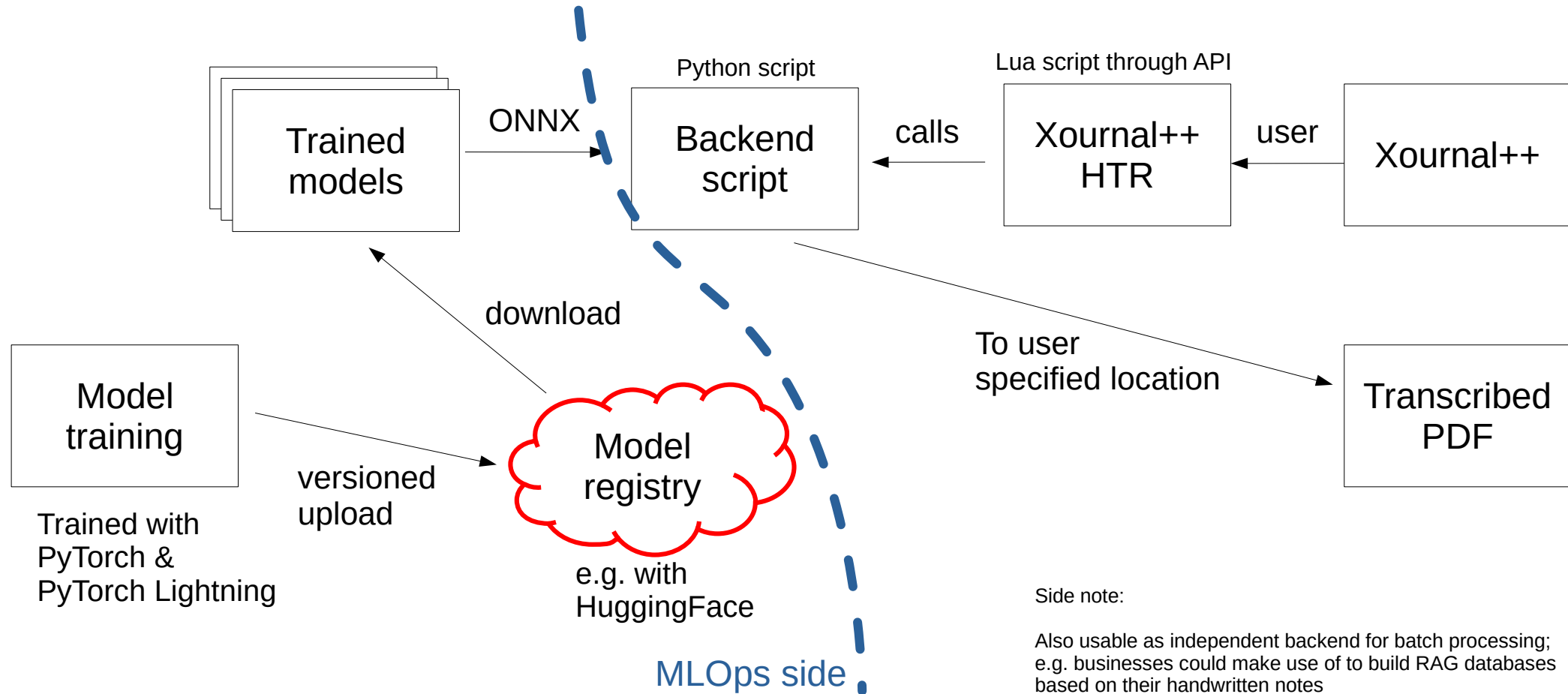
Xournal++ HTR is available online



All this already public!

https://github.com/PelleNitram/xournalpp_htr

System design



Excursus: Offline vs Online HTR

Image data



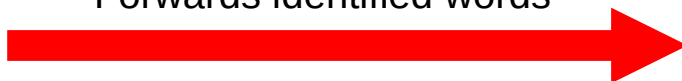
Current ML model

Zero shot: Offline HTR with HTRPipeline by Harald Scheidl

Step 1:

Recognition

Forwards identified words



Step 2:

Transcription

Segmentation:

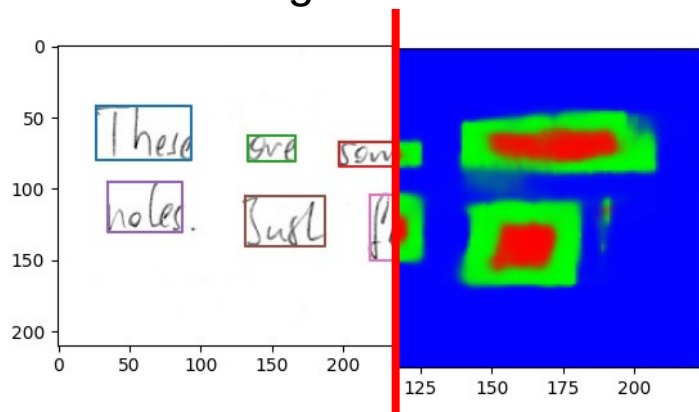


Image to sequence:

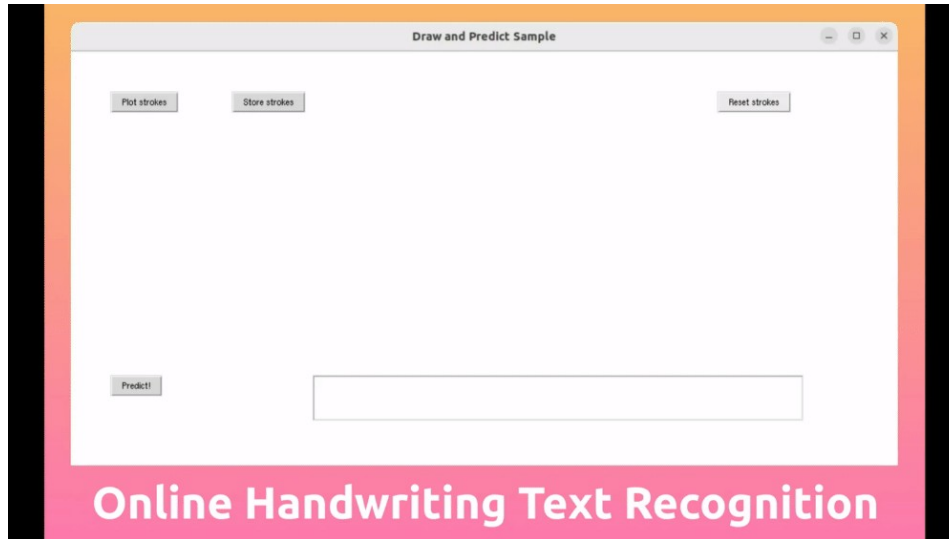
word → "word"

Sources:

- <https://github.com/githubharald/WordDetectorNN>
- <https://github.com/githubharald/SimpleHTR>
- <https://github.com/githubharald/HTRPipeline>

Next ML model

I already developed a model to perform online HTR text recognition:



Uses:

- LSTM layers
- CTC loss
- CTC decoding

Weights: 5MB in size

Available at: <https://github.com/PelleNitram/OnlineHTR>

Conclusions

- We live in a world in which ML features now come to FOSS applications
 - Example: I built **Xournal++ HTR** which will impact **>10k users**
- This requires MLOps practises ...
- ... that might differ from cloud-native MLOps b/c ML runs on the edge (i.e. on the user's local device)

Go out and build awesome ML open source features!

How to select project to contribute?

Actionable advice

- Active
- Be proactive in both ideating and implementing a PoC
- Plugin system is useful
- Language you know
- Available zero shot model is useful
- Get in touch with community early on

Thanks!

Do you have any questions?



Martin Lellep