# PART I: Analysis on Single Non-Recursive Models

**Evolution of the accuracy**A close up of a map

Description automatically generated**Evolution of the loss:**A screenshot of a cell phone

Description automatically generated

Based on the evolution of the loss, it looks like adding more filters helps achieve better results (as opposed to LeCun conclusions) but at the price of overfitting. The fact that the loss increases at the same time as the accuracy could be that the correct class is still being identified, but the certain put by the network on that class is lower, and the scores on the prediction are softer.

On the other hand, with a constant number of filters, it looks like having the double of layers is not helping much in reducing the error. Also, it looks like the depth is not responsible for the overfitting.A close up of a logo

Description automatically generated

**INFERENCE TIME**

~~A close up of a logo

Description automatically generated~~

~~A screenshot of a cell phone

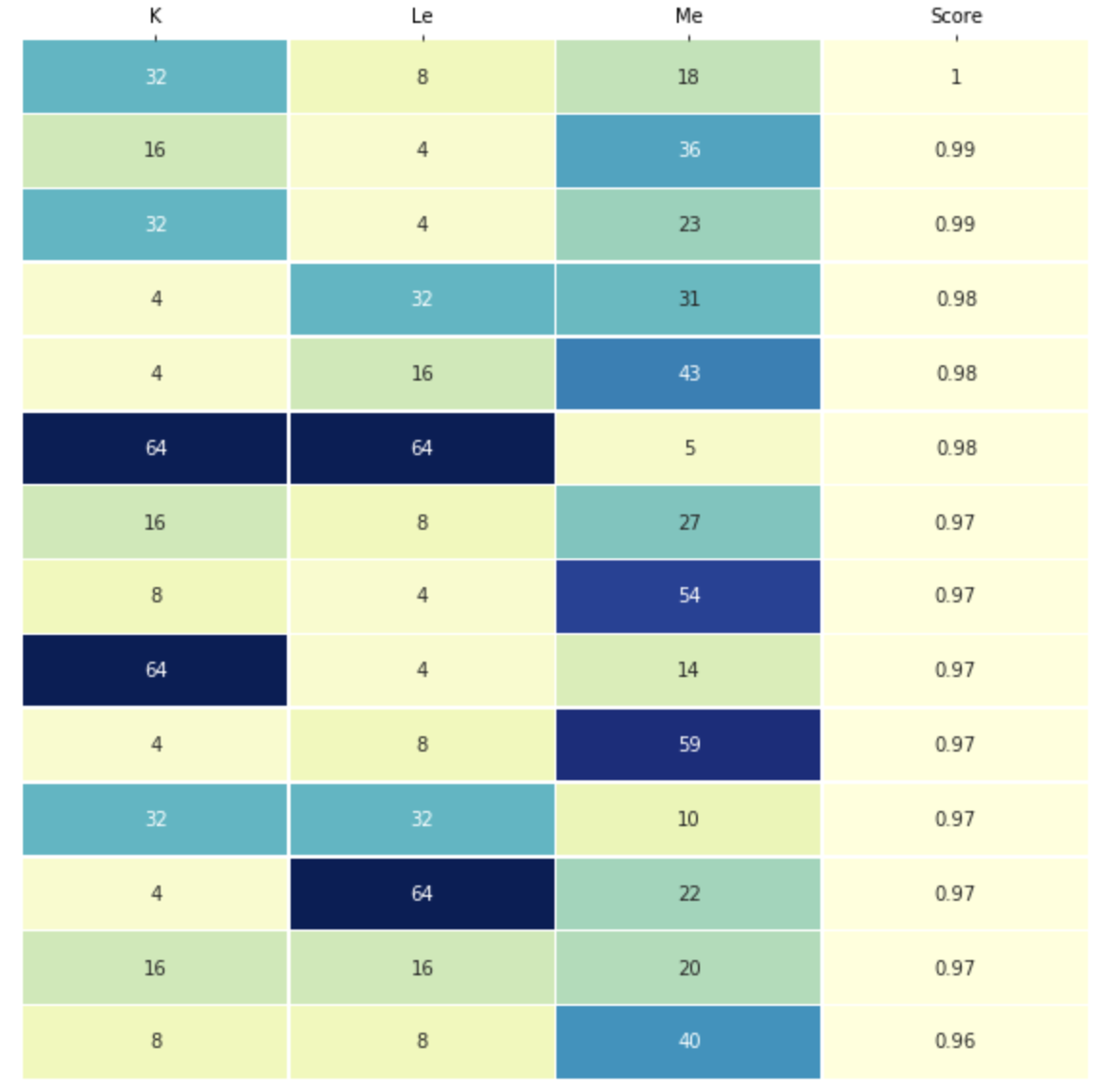
Description automatically generated~~

# A close up of text on a white background Description automatically generated

# PART II: Looking for candidates to build ensembles

Choose a single deep networks --> L = 32, M = 64

Run deep\_to\_ensemble:



Look for diverse candidates among these with reasonable numbers:

|  |  |  |
| --- | --- | --- |
| **Le** | **Me** | **K** |
| 4 - Very low | 36 - Large | 16 - Large |
| 16 - Large | 31 - Large | 4 - Low |
| 32 - Very large | 10 - Low | 32 - Very large |
| 8 - Low | 59 - Very large | 4 - Low |
| 16 - Large | 20 - Low | 16 - Large |

It could be interesting to assess with these candidates the impact of K.

However, we need to discuss if L=32 and M=64 is a good starting point.