Task 8 : Survey on NAND Flash Memories and Classic Machine Learning Techniques

Sanjeev Narayanan

30th March 2018

1 NAND Flash Memory

NAND flash memory is a type of nonvolatile storage technology that does not require power to retain data. An important goal of NAND flash development has been to reduce the cost per bit and to increase maximum chip capacity so that flash memory can compete with magnetic storage devices, such as hard disks. NAND flash has found a market in devices to which large files are frequently uploaded and replaced. MP3 players, digital cameras and USB flash drives use NAND technology.

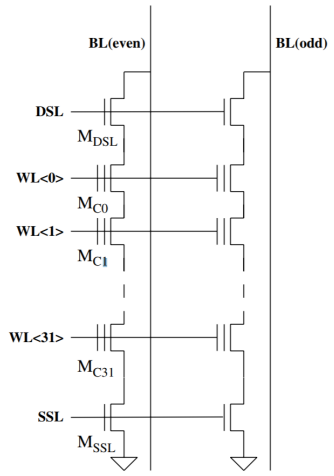

Figure 1: Matrix Structure in NAND Architecture

In the NAND Flash architecture, the cells are connected in series, in groups of 16 or 32. Two selection transistors are placed at the edges of the stack, to ensure the connections to ground. The basic structure is shown in Fig 1. When a cell is read, its gate is set to 0 V,while the other gates of the stack are biased with a high voltage (typically 4–5 V),so that

they work as pass-transistor, regardless of their threshold voltage. An erased NAND Flash cell has a negative threshold voltage; on the contrary, a programmed cell has a positive threshold voltage but, in any case, less than 4 V. In practice, driving the selected gate with 0 V, the series of all the cells will sink current if the addressed cell is erased, otherwise no current is sunk if the cell is programmed.

The reading method in NAND memories is the charge integration, which uses the parasitic capacity of the bitline. This capacitance is precharged to a fixed value (typically 1.2 V): if the cell is erased, it sinks current and discharges the bitline; otherwise, if it is programmed, it does not sink current and the bitline keeps its initial value.

The "program operation" is the writing of the information in a memory cell and it is usually performed by transferring the electrons from the substrate of the cell into its floating gate; in this way the threshold voltage of the cell is increased. In a NAND flash memory cell, this is achieved by s exploiting the quantum effects of tunnelling of electrons in the presence of a high electric field. In particular, the operation depends on the polarity of the electric field: if it is directed from substrate to gate, than a program of the cell is obtained; an erase operation occurs if the polarity of the electric field is the opposite one [1].

Machine learning can be used in flash memory devices to optimize storage capacity, endurance and data retention. The way that data is stored in flash memory devices leads to wear and limitations in data retention with repeated erase and program cycles and even with multiple reading of cells. These issues become even more important as the number of bits per flash memory cell increases. The tuning of registers in 2D as well as 3D flash memory can be used to optimize flash wear, storage capacity and data retention; but as the number of these registers increases, this task becomes impossible to do manually.

2 Basic Machine Learning Techniques

2.1 Supervised Learning

The supervised machine learning algorithms are those algorithms which needs external assistance. The input dataset is divided into train and test dataset. The train dataset has output variables which needs to be predicted or classified.

2.1.1 Decision Trees

Decision trees are those type of trees which groups attributes by sorting them based on their values. Decision tree is used mainly for classification purpose. Each tree consists of nodes and branches. Each nodes represents attributes in a group that is to be classified and each branch represents a value that the node can take

2.1.2 Bayes Classifier

Naïve Bayes mainly targets the text classification industry. It is mainly used for clustering and classification purpose. The underlying architecture of Naïve Bayes depends on the conditional probability.

2.1.3 Support Vector Machine

Another most widely used state-of-the-art machine learning technique is Support Vector Machine (SVM). It is mainly used for classification. SVM works on the principle of margin calculation. It basically, draw margins between the classes. The margins are drawn in such a fashion that the distance between the margin and the classes is maximum and hence, minimizing the classification error. [2][3]

2.2 Unsupervised Learning

The unsupervised learning algorithms learns few features from the data. When new data is introduced, it uses the previously learned features to recognize the class of the data. It is mainly used for clustering and feature reduction.

2.2.1 K-means Clustering

Clustering or grouping is a type of unsupervised learning technique that when initiates, creates groups automatically. The items which possesses similar characteristics are put in the same cluster. This algorithm is called k-means because it creates k distinct clusters. The mean of the values in a particular cluster is the center of that cluster.

2.2.2 Principal Component Analysis

In Principal Component Analysis or PCA, the dimension of the data is reduced to make the computations faster and easier. [2][3]

2.3 Neural Networks

The neural network (or artificial neural network or ANN) is derived from the biological concept of neurons.It works on three layers. The input layer takes input (much like dendrites). The hidden layer processes the input (like soma and axon). Finally, the output layer sends the calculated output (like dendrite terminals) [2][3]

References

[1] L Crippa, R Micheloni, I Motta, and M Sangalli. Nonvolatile memories: Nor vs. nand architectures. In *Memories in Wireless Systems*, pages 29–53. Springer, 2008.

[2] Ayon Dey. Machine learning algorithms: A review. *IJCSIT) International Journal of Computer Science and Information Technologies*, 7(3):1174–1179, 2016.

[3] Aurélien Géron. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems.* " O'Reilly Media, Inc.", 2017.