

AS227450

Bimhox: The evolutionary information gene

Alberto Tono
HOK

Learning Objectives

- Understand a Paradigm Shift
- Evolutionary Information
- Practical Use case: Tensorflow

Description

This class will talk about the DNA of the different buildings, how their can be compared to living organisms. We can classify species so why we can't transfer the knowledge from Bioinformatic and apply to the AEC industry?

This class will show how Deep learning models can be applied to perform predictive analysis based the "parents" generation.

The concept derive from a better understanding of our human evolution works. It implies an evolution of our genes during the transition of the information from the parents to the children, from previous airports to the next one. The genes that allow this is called HOX. Understanding the evolution theory related to the transfer knowledge is a key point for the raising of the Artificial Intelligence in the AEC industry in order to make our information evolve with the previous lessons learnt provided by the constructive design ideas.



Speaker



Alberto Tono serves as Research and Development at HOK. Prior to HOK, Alberto is working in the computational design space since 2015 for BIM workflow on large projects. His studies took him to Harbin Institute of Technology (HIT) in China and to Rene van Zuuk Architect Studio in the Netherlands. He is very keen on computational design and machine learning. Learning and this passion inspired him to found and helped various community ,Deeplearningitalia, Italian Dynamo User Group, San Francisco Computational Design User Group, Bimhox. He also is speaker in several international conferences.



Contents

Bimhox: The evolutionary information gene	1
Tackling problems:	4
Approach to Technologies:	5
Immutable information.....	5
Classification.....	6
Prediction.....	7
Biomimicry approach	10
Historical Background.....	10
Evolutionary theories	10
Hox background.....	11
Bioinformatics and Transfer Learning	12
The Framework.....	15
References	18

Tackling problems:

“The information is static!”

What kind of problems you can solve with Artificial Intelligence, and how can you apply it in your BIM workflow. This paper took inspiration from nature, from hox gene and cluster¹ in order to combine the bioinformatics knowledge and merge then in the Building Information Modeling (BIM) ecosystem to develop a higher goal: Building Knowledge Modeling (BKM)². Also you can understand what **opportunities will you lose** if you won't apply this in your business.

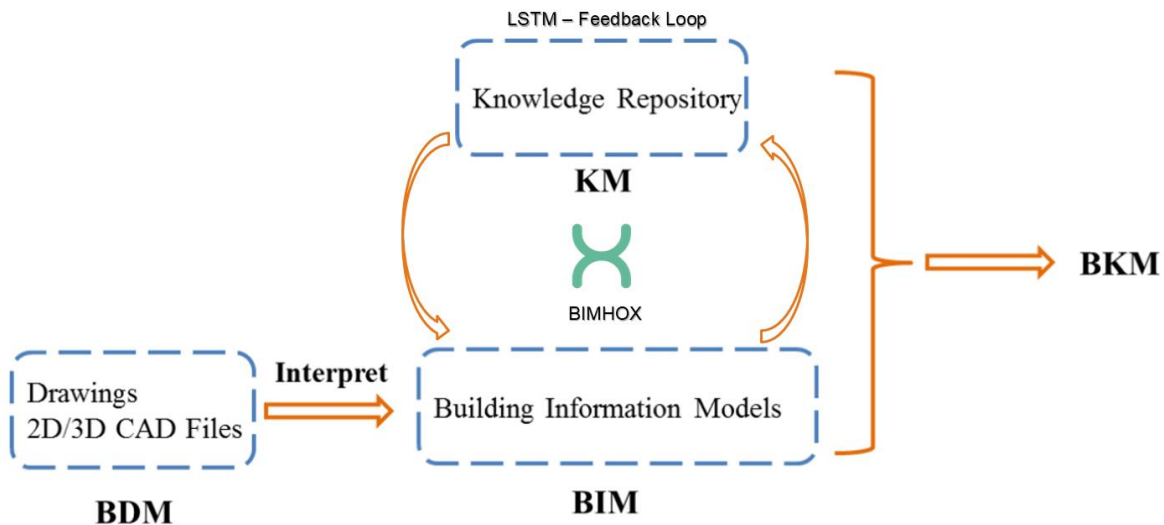


Figure 1 Development of BDM, BIM and BKM with a continuous feedback loop

After all the theoretical conversation about the implementation of Knowledge in the AEC from King, Davenport, Prusak Bhatt it has been decided to approach the problem from a computer science background with a more practical approach. The implementation of a knowledge management should always enable capturing and storing knowledge in an efficient way (Kamara 2002) but most importantly this information needs to evolve during the process.

¹ Hox Genes in Development: The Hox Code, By: PZ Myers, Ph.D. (University of Minnesota, Morris) © 2008 Nature Education

² From Building Information Model to Building Knowledge Model, Renate Fruchter, Tim Schrotenboer Gregory P.Luth, 2009. Conference. International Workshop on Computing in Civil Engineering

Approach to Technologies:

Technologies are used to enhance our workflow, processes and to solve problem related to them. Sometimes, tackling a new problem required a different point of view and a different question.

As Albert Einstein said, “We cannot solve our **problems** with the same thinking we used when we created them.”³ this require a Khunian approach to better different issues. As It has been anticipated in the abstract, this paper is focused on transfer learning from different fields like Bioinformatics and Computer Science to the AEC, BIM fields. In this way it will be possible to tackle some of the problems related to the AEC with a totally different point of view.

One of the most important cutting edge-technology that it will be analyzed in this paper it is deep learning. As you may noticed it has been produce multiple research papers without a clear explanation of the real process under the hood for some artificial intelligence application, but an infinite trial and errors process to increase the accuracy of the models. This paper will be focus on Explainable AI whose actions can be understood by human and explained accordingly also with the GDPR as well.

The approach of computer science researchers during they trail error process adopted a totally different approach compared to the one that the AEC industry in general seems to follow. The AECO is strongly related to **standard**⁴ and procedures. This approach require a big Paradim Shift in the way in which we are used to work. A great example of how our industry should work is the Transbay Center in San Francisco. In this project the water management system was totally illegal accordingly with the California Standards, but they were able to prove the value of their design and now it is mandatory for the entire state.

After we will overpass this old paradigm in order to approach the new one.

What are some of the main problems that will be analyzed in this paper:

- Immutable information
- User Classification
- Unpredictable future

Immutable information

The **meaningful information in our design processes does not evolve** by itself. They are allocated in the designer’s mind and they create his/her knowledge and wisdom of the process. This knowledge is the final process of a series of mistakes, lessons learn, design changing and an endless series of variable. If we think about the AEC final outputs, these are composed by printed sheets or better PDFs that are considered as the “black holes” for the information. How to make sure that our **previous projects** will be able to inform our decision and **future design outcome** avoiding repetitive errors and waste of time embedded in previous generation. Furthermore, what does it means to allow information to evolve and what kind of advantages we

³ Complex Problem Solving and Its Position in the Wider Realm of the Human Intellect, Samuel Greiff, and Ronny Scherer.

⁴ Building Information Modeling: Analyzing Noteworthy Publications of Eight Countries Using a Knowledge Content Taxonomy, Bilal Succar, Mohamad Kasseem, Nashwan Dawood. 2015.

can derive from this evolution and from the control and monitoring of its process. Like evolution in Nature.

Classification

Monitoring and controlling the process in which people collaborate, work it is crucial to better understand the point of failure and implement lean system on top. Understand the level of skills and the quality of the work done without invading the privacy of the designers and measure their performance it allows to avoid waste in the process. In enterprises designers spend more than 70% of their time working with BIM Software. Also, almost 15% of the resource are allocated for training and improvement of software and designer skills. One of the products that we developed in order to collect data is called Mission Control, but from there to understand based on this data is a user is good or not it required a lot of time and a constant monitoring of the operation that he or she perform. How train the right people and in what is always more difficult, but also how to assist them in a better way in the design process. Artificial Intelligence can solve this issue as well. While we have a unique Taxonomy for the Biological world we have an infinite type of classification for the AEC industry: Unifomat, Omniclass, etc,⁵ and in BIM, we have the Element division of different file format as Revit, IFC and so on and so forth. But instead of focusing in classify the building why we cannot classify the users, and put the Human at the center of the AI first step. Start training our machine in order to classify which user is performing better in a particular building typology but also be able to understand good and bad users and be able to automatically target specific people with a specific training required. In this way we can provide small **reward** to the good users, who respect the company standard, safe time and help each other to make them more engage in the workflow while we are constantly **monitoring** their performance with a careful attention in avoiding possible bias.⁶

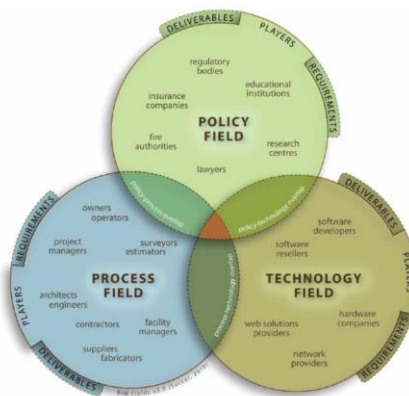


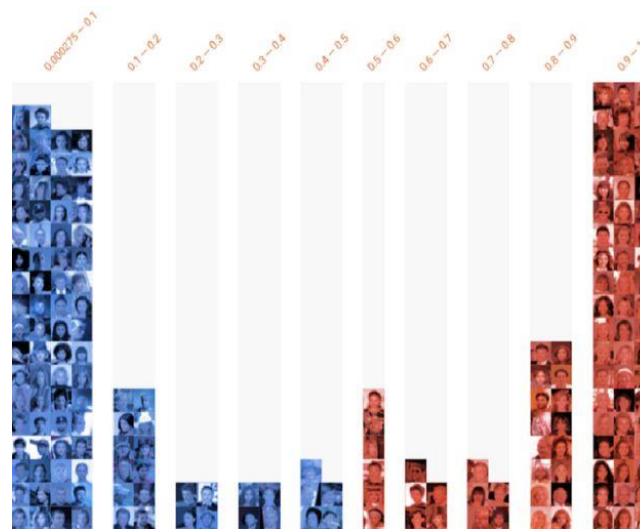
Figure 2 The interlocking fields of BIM activity, special focus after the GDPR in order to carefully evaluate these solutions.

The limitation in not having a dynamic and evolving flow of distilled information capable to classify the users don't allow to foresee a possible problem because it required a deep analysis of the data and an incredible amount of time if this is not automatic. Furthermore, it will be difficult to

⁵ BSA. "Frequently Asked Questions About the National BIM Standard-United States™". <http://www.buildingsmartalliance.org/index.php/nbims/faq/> (retrieved on Nov., 2012)

⁶ Measuring BIM performance: Five metrics Bilal Succar, Willy Sher and Anthony Williams

allow third party software to acquire that knowledge developed in years of practice and sacrifice. While Nature is doing all this work since 3.8 billion years ago we have the enormous advance that we can start right now with its experience on our shoulders, thanks to AI capabilities these tasks will be achievable.



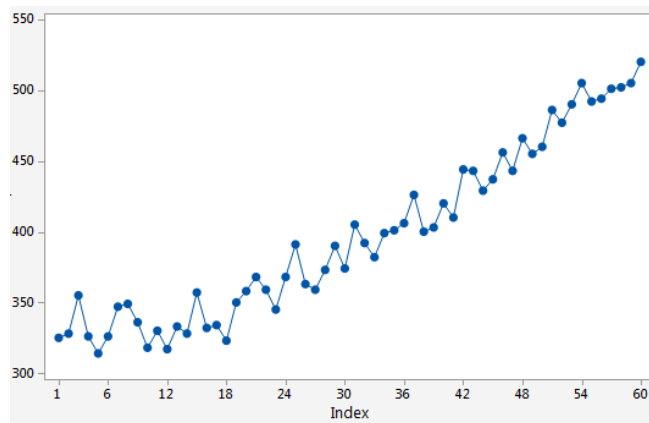
[Image Graph of Classification/ Bias detector: Age, Gender etc]

Prediction

In the AEC it is more complex than in finance or other fields to make prediction. This because the AEC is lacking a data-driven approach that could be more practical than the financial application, in the stock market these model are applied even if a blindfolded monkey throwing darts at a newspaper stock listing should do as well as any investment professional. For example in the bioinformatics, nowadays we are aware that, during the transmission of traits from parents to offspring, the information evolve and allow an **“intelligence” evolution** based on the hox that facilitate the correct transcription findings that correlation with interesting morphological changes in evolution. This intelligence came from a sort of prediction of the future consequences based on the previous experiences, based on all the previous lesson learnt. This feature was one of the first when the main concept of Machine Learning started in the 1959 by Arthur Samuel with the idea of Predictive Analytics and the idea to program machine to “learn” from previous data like our DNA was doing for 3.8 Billion of years. Forecasts are important in order to better understand how well our design is performing, having a digital assistant during the design process will prevent mistakes and issues. As James Watson “the essence of life is information” that it is why we started from a data-driven design and now passing thought a building information modeling one we are heading to a design with knowledge. (DKIW Pyramids).



Figure 3 DIKW Piramid



[Graph of Prediction and Time series]

LSTM Long-Short Term Memory⁷ that is a time-series model in order to make predictive analysis, these models can look at the history of a sequence of data and correctly predict what the future elements of the sequence are going to be. In this way the action of our designers and engineers can be track and their lesson learned can be memorized and used as a support for their decision during their workflow.

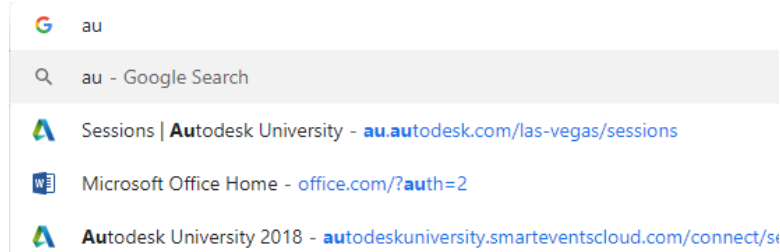


Figure 4 Prediction such as the one done by Google

⁷ <https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/>, Long-Short Term Memory LSTM Hochreiter 1997. RNN, Memory Management

It will be always important to collect all the data with their time-stamp in order to have a specific understanding of the process and of their sequence. Imagine to keep control of the file size of the model and correlate these data with the review with the client and also with their request, you can keep track of the time during the process when usually the model has it major impact in the production and also when it will required a BIM Audit. Specifically we should use a momentum-based algorithm because they are not focus in the number but in the vector, if the file size is going up or down. This could be really useful in order to exploit hidden pattern in the data and to perform useful prediction for the QAQC period.

Biomimicry approach

Nature have inspired engineers and problem solvers since 1M years ago with fire control system, but nowadays it has even more important than ever, it touches almost all the new discoveries because we have a deeper understanding about how it works. Some great examples are provided by the works from biomimetics, biomimicry, and other disciplines⁸. Otto Schmitt, Frei Otto, Leonardo da Vinci Buckminster Fuller. Nature can inspire also different fields such as Building information modeling, architecture, engineering and a lot of others, for instance, evolutionary computation and multi-objective optimization (MOO)⁹ are being used to allow facilitate and empowered the design of some of the modern buildings. However, none of these solutions and methods really embraced the power and potential provided by the AI capabilities and Neural Networks, apparently another discovery that try to mimic how the human brain works.

In 2000, the conversation about the interaction between design and the natural world asked, “How can innovation in the built environment be inspired by nature?” One interesting aspect to the biomimetic approach is the focus on the information and knowledge of our BIM models. There is 3.8 billion years of natural world around us. What can we learn about these processes and how can we translate that into our workflow and integrate in a holistic environment driven by an Artificial Intelligence? Indeed, it is, with the lessons learnt from the past as Lamarck, Darwin, James Watson and many other biologists, bioinformatics, genetists, that we aim to address these problems

Historical Background

First of all it is important to analyze some of the different evolutionary theories to understand how this conclusion is relevant to develop a framework for artificial intelligence in the design processes, or better we should called at this point as “Natural intelligence”¹⁰.

Evolutionary theories

As Buckminster Fuller said “We do not seek to imitate nature, but rather to find the principles she uses.” Moreover we started to study the process of the evolution in the living being. It is well known that “God doesn’t play dice with the universe”, randomness it only a quick solution to something that could be attributed to human ignorance about the details of the processes involved. But Nature in its evolution and structure has its own intelligence¹¹.

At the begging of the 18th century the *Lamarckism* theory took place. It is the hypothesis that an organism can pass on characteristics that it has acquired through use or disuse its lifetime to its offspring being part of the Theory of Inheritance. In an AEC environment this could be compared to a *restauration* of a building that adapts itself to the need of the users and hands down some cultural values.

Otherwise, few decades later *Darwin* formulated his evolution theory in the *On the Origin of Species* summarize like to natural selection where the fittest survive. In comparison to a building that after a lot of year remain a symbol or a landmark for a specific place.

⁸ Biomimicry Innovation inspired by Nature Janine Benyus, Janine M Benyus 2002.

⁹ Multi-Objective Optimization in Architecture, ian Keogh and David Benjamin 2010.

¹⁰ J. Weng, Natural and Artificial Intelligence: Introduction to Computational Brain-Mind, BMIPress, 2012.

¹¹ Brain S. Everitt, Chance Rules (1999), Chapter 12, p. 175.

Nowadays there is a particular hype around the *Hox Cluster* that is going to redefine the way in which we thought about evolution. This series of genes that will explain later in details embraces the theory that our evolution has a sort intelligence in itself and it is not totally random. It will be like looking a building like a summary of all the historical experiences from the past.

Hox background

The story of the Hox gene started with the pioneering genetic studies of Edward B. Lewis at Caltech, Nobel Prize for the Medicine. Lewis found out that the normal function of these homeotic genes is to assign distinct or positional identities to cells in different regions. **Such as the architects that need to manage the space program.**

Mario Capecchi, another Nobel Prize, determined that the placement of cellular development in the proper order along the axis of the body from head to toe. This started to process the information that there were a pattern that can define the way in which organism are build. Practically defined that the generation of an organism has a specific order. The **Hox cluster protein expression pattern**¹² has long been studied for its ability to determine the body plan of *metazoan organisms* for decades, in terms of it is able to determine what a cell will begin differentiating towards its parts such as arm, leg, brain, and etc. Recently it is possible to measure the physical looping structure of DNA at unrepresented resolution using a set of techniques based around: Chromosome conformation capture (3-C). It is used to analyze the spatial organization of chromatic in a cell.

The Hox cluster actually expresses from left to right according to the location of the tissue, from left to right, anterior (head) to posterior (butt) axis (in the simpler metazoans at least):

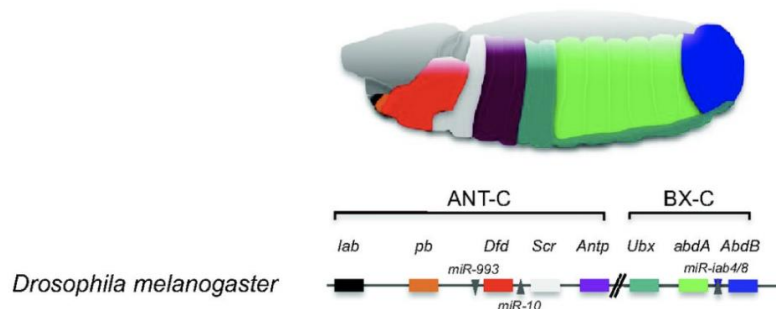


Figure 5 Left = anterior, right = posterior

The way in which we build building has more often followed a specific order starting from the foundation to the roof but more important the way in which architect solve test fit issues can be related and can learn from how the 3-C does.¹³

¹² Duboule, D. (1992). The vertebrate limb, temporal colinearity and the Hox/HOM gene network. BioEssays 14, 375-384

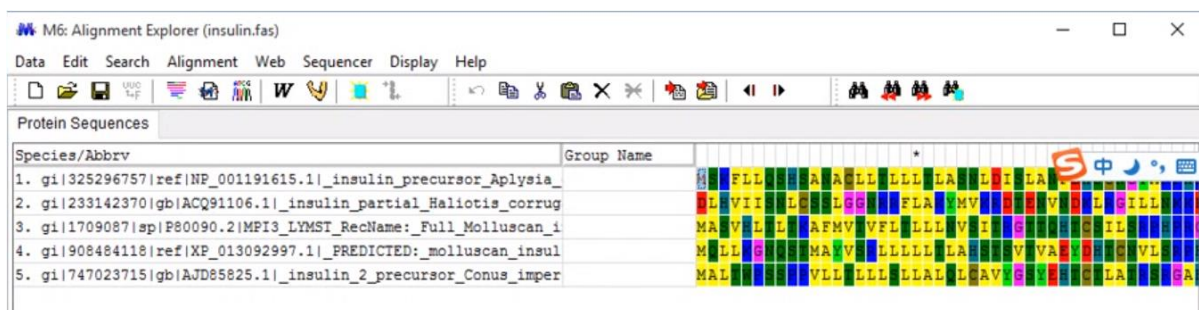
¹³ (Hagfish and lamprey Hox genes reveal conservation of temporal colinearity in vertebrates <https://www.nature.com/articles/s41559-018-0526-2>, Juan Pascual-Anaya, Iori Sato, Fumiaki Sugahara, Shinnosuke Higuchi, Jordi Paps, Yandong Ren, Wataru Takagi, Adrián Ruiz-Villalba, Kinya G. Ota, Wen Wang & Shigeru Kuratani)

T, C and G, which in succession, in a precise order, create the genetic code of the individual. 3 billion of these letters are needed to build the entire genome of a human being. Understanding and simplifying the biological ideas can facilitate the machine to accomplish complex and unthinkable tasks with the today technologies, such as predicting how it looks like someone by his/her genetic code. It can be compared to the way in which architect should start their design, looking at the data in order to envision how the design will look like.

Nowadays, buildings are represented as a **digital relational database of information**. We consider a building not as a simple sum of areas, technical elements, construction materials and plants, but as an organic system in which each element relates to each others in a complex way in order to satisfy the needs of the users. In the standard design two main categories of parameters are always taken into account: those relating to the characteristics of the spaces (environmental system) and those relating to the characteristics of the elements within the spaces (technological system). Information in **BIM** technology plays a fundamental role, they are integrated and structured within 3D models. In analogy to genetics where the genes are building's DNA molecules. We can define species and we can **classify** building typologies such as.

To classify building from BIM models it is important to deeply understand how the information are structured and their differences peculiarities. Why don't "train" a machine in order to do that. This can be a significant small steps in order to achieve other great results.

In a dazzling and still mysterious feat of self-organization, these cells arrange themselves into a complete organism, in which bone, muscle, brain and skin integrate into a harmonious whole.¹⁴ Some information of the building element should reside on it for the whole lifecycle of the building organism in order to predict and monitoring their behaviors. Also it could be possible to simplify the building as a sequence of bases (such as ACGT) for example the entrance could be E, the office space could be O and so on. After it will be possible to describe the building as a sequence of bases when can use the knowledge from the Biology. In biology are used software such as: BLAST, CRISPR¹⁵ and alignment score and MEGA 6 to sequence protein to predict possible outcomes and classify DNA strings. In order to train the machine to discover pattern and insight. In this way could be easier to classify buildings and also make prediction based on this theory.



At the same way we could classify buildings but even more the users of these technologies. We could classify also the Users and their behavior in the modeling and design process, avoiding

¹⁴ From The Molecular Architects of Body Design. William McGinnis and Michael Kuziora Scientific American February 1994.

¹⁵ <https://blast.ncbi.nlm.nih.gov/Blast.cgi/> / http://sbml.org/Main_Page

bias given by age, sex, background, education and others putting always human at the center of the design processes in order to empower the outcome and track reliability. Some solutions that right now provide a quick way to use AI in business are provided by Oracle, IBM, Google, AWS but also other companies such as Gamma, Clarifai. These solutions offer a deep data analytic way to analysis data. Web dashboard are becoming more and more important: the efforts provided by Unify, Clarity and our Mission Control (HOK) are a clear examples that demonstrate the need of a better data analytic and visualization.

The Framework

How technically it will be possible to solve all the previous issue and address the solutions mention above. For this experiment it has been used Tensorflow and Keras. There are also some future idea to implement with Tensorflow.js in order to run directly with Forge on the website¹⁶.

Tensors are immutable in tensorflow, you can only generate operation in order to create new tensors.

Variables are used to store data and then reassign the updated value during the model training.

Ops allow to manipulate the data in order to produce new tensors, a new children

Genes that evolves Building A...HOX A Building B...HOX B HOX (H=0 - O=1 - X=2)

- 1) When architects approach a new project they try to find inspiration from art or from the surrounding building, local tradition but when they perfume these researches they usually start from the zero like a writers with his blank page at the begging of theirs novels

```
1 const conceptualdesign = tf.zeros([5]);
2 const biases = tf.variable(conceptualdesign); // initialize biases
3 biases.print(); // output: [0, 0, 0, 0, 0]
```

```
const conceptualdesign = tf.zeros([5]);
const biases = tf.variable(conceptualdesign); // initialize biases
biases.print(); // output: [0, 0, 0, 0, 0]
```

- 2) This value can get update after we receive the data that can drive our design.

```
1 const updatedValues = tf.tensor1d([0, 1, 0, 2, 1]);
2 biases.assign(updatedValues); // update values of biases
3 biases.print(); // output: [0, 1, 0, 2, 1]
```

```
const updatedValues = tf.tensor1d([0, 1, 0, 2, 1]);
biases.assign(updatedValues); // update values of biases
biases.print(); // output: [0, 1, 0, 2, 1]
```

- 2) How better can be instead if we sum the previous information of other building in order to produce a new one that has embedded in itself the data of the previous one

```
1 const buildingA = tf.tensor2d([[1.0, 2.0], [2.0, 0.0]]);
2 const buildingB = tf.tensor2d([[0.0, 1.0], [1.0, 2.0]]);
3
4 const buildingA_plus_buildingB = buildingA.add(buildingB);
5 buildingA_plus_buildingB.print();
6 // Output: [[1 , 3 ],
7 //          [3, 2]]
```

¹⁶ For example analyze the user facial expression and perform specific operation accordingly with their behavior.

```
const buildingA = tf.tensor2d([[1.0, 2.0], [2.0, 0.0]]);
const buildingB = tf.tensor2d([[0.0, 1.0], [1.0, 2.0]]);

const buildingA_plus_buildingB = buildingA.add(buildingB);
buildingA_plus_buildingB.print();
// Output: [[1 , 3 ],
//          [3, 2]]
```

- 3) The key property of tensorflow is to have a chainable API in order to allow us to call ops on the results of ops.

```
1 const sq_sum = buildingA.add(buildingB).square();
2 sq_sum.print();
3 // Output: [[1 , 9 ],
4 //          [9, 4]]
```

```
const sq_sum = buildingA.add(buildingB).square();
sq_sum.print();
// Output: [[1 , 9 ],
//          [9, 4]]
```

Models

The models are our HOX gene that given some information from our parents it will produce an output. In order to approach the problem with a high-level API it is really important to construct our models out of layers, key concept for deep learning.

Easiest, works when the model is a simple **stack** of each layer's input resting on the top of the previous layer's output. But Models are a lot more flexible in term of organization they can be subjected to branching or skipping.

```
1 const model = tf.sequential();
2 model.add(
3 //First layer must have an input shape defined
4   tf.layers.simpleRNN({
5     units: 20,
6     recurrentInitializer: 'GlorotNormal',
7     inputShape: [80, 4]
8   })
9 );
10
11 const optimizer = tf.train.sgd(LEARNING_RATE);
12 model.compile({optimizer, loss: 'categoricalCrossentropy'});
13 model.fit({x: data, y: labels});

const model = tf.sequential();
model.add(
```



```
//First layer must have an input shape defined
tf.layers.simpleRNN({
  units: 20,
  recurrentInitializer: 'GlorotNormal',
  inputShape: [80, 4]
})
);

const optimizer = tf.train.sgd(LEARNING_RATE);
model.compile({optimizer, loss: 'categoricalCrossentropy'});
model.fit({x: data, y: labels});
```

Conclusion

Bimhox want to be a safe place for researchers that aim to connect and transfer their knowledge in different fields. This is an example about how Bioinformatic can inspire BIM. The hope is that the information in our model need to evolve somehow in order to achieve the knowledge level.

References

- Aamodt, A., & Nygård, M. (1995). "Different roles and mutual dependencies of data, information, and knowledge-an AI perspective on their integration". *Data & Knowledge Engineering*, 16, 191-222.
- Ackoff, R.L. (1990). "From data to wisdom". *Informatie*, 32, 486-90.
- Bellinger, G., Castro, D., & Mills, A. (2004). "Data, information, knowledge, and wisdom". <http://www.systemsthinking.org/dikw/dikw.htm> (retrieved on Dec., 2012).
- Bhatt, G.D. (2001). "Knowledge management in organizations: examining the interaction between technologies, techniques, and people". *Journal of Knowledge Management*, 5, 68-75.x
- Chen, M., Ebert, D., Hagen, H., Laramée, R.S., Van Liere, R., Ma, K.-L., Ribarsky, W., Scheuermann, G., & Silver, D. (2009). "Data, information, and knowledge in visualization". *IEEE Computer Graphics and Applications*, 29, 12 -19.
- Davenport, T.H., & Prusak, L. (2000). "Working knowledge: how organizations manage what they know". Harvard Business Press, 1998.
- Fruchter, R., Schrottenboer, T., & Luth, G.P. (2009). "From Building Information Model to Building Knowledge Model". *Proceedings of the 2009 ASCE International Workshop on Computing in Civil Engineering*, Austin, TX, USA, 380-389
- Kamara, J.M., Augenbroe, G., Anumba, C.J., & Carrillo, P.M. (2002) "Knowledge management in the architecture, engineering and construction industry". *Construction Innovation: Information, Process, Management*, 2, 53-67.
- King, W.R. (2007). "Knowledge management: a systems perspective". *International Journal of Business and Systems Research*, 1, 5-28.
- Kock, N.F., McQueen, R.J., & Corner, J.L. (1997). "The nature of data, information and knowledge exchanges in business processes: implications for process improvement and organizational learning". *The Learning Organization*, 4, 70-80.
- Liu, F., Jallow, A., Anumba, C., & Messner, J. (2013). "Integration of change and knowledge management processes in energy efficient retrofit projects". *Proceedings of 2013 ASCE International Workshop on Computing in Civil Engineering*, Los Angeles, CA, USA.
- Ribino, P., Oliveri, A., Re, G.L., & Gaglio, S. (2009). "A knowledge management system based on ontologies". *Proceedings of the 2009 International Conference on New Trends in Information and Service Science*, Washington, D.C., USA, 1025-1033.
- Duboule, D. (1992). The vertebrate limb, temporal colinearity and the Hox/HOM gene network. *BioEssays* 14, 375-384.
- Chun, M., Sohn, K., Arling, P., & Granados, N.F. (2008). Systems theory and knowledge management systems: The case of Pratt-Whitney Rocketdyne .Paper presented at the *Proceedings of the 41st Hawaii InternationalConference on System Sciences*, Hawaii.
- Ranzato, M., Boureau, Y., & LeCun, Y. (2008). Sparse feature learning for deep belief networks. *Adv. Neural Inf. Proc. Sys.* 20 (pp. 1185–1192).