



UNIVERSITAT DE  
BARCELONA

Treball final de grau

GRAU EN INFORMÀTICA

Facultat de Matemàtiques i Informàtica  
Universitat de Barcelona

---

**MACHINE LEARNING  
APLICAT A LA GENERACIÓ  
DE MOODBOARDS**

---

**Autor: Ferrer Margarit, Marc**

**Director: David Buchaca Prats**  
**Realitzat a: Interiorvista**

**Barcelona, 27 de juny de 2018**

## Abstract

Moodboards are used to represent objects that are formed by the composition of other objects and they try to obtain a result of cohesion among them and try to generate a new image. Machine learning allows us to train a machine so that it can generate the values with the data that we have entered previously, which must be correct data.

What we want to obtain is a combination between the two parts to achieve the generation of new images, with correct and cohesive objects and different from the images used to train the machine.

So the objective of this project is to create a machine that allows us to create moodboards and generate correct it without the need of a person to create them, and design an automated system to replace a slowly non automated task.

Apart from creating this system, we will also analyze it's efficiency, between the non-automated system and the automated system, so we can establish whether the system created is efficient enough, fast and accurate generating moodboards.

Finally, these tests will be carried out by supervising expert people in order to improve the evaluation.

## Resum

Actualment els moodboards s'utilitzen per representar objectes que estan formats per la composició d'altres objectes i pretenen obtenir un resultat de cohesió entre ells i generar una imatge nova. El machine learning ens permet entrenar una màquina per tal que aquesta sigui capaç de generar els valors amb les dades que li hem entrat prèviament, les quals han de ser dades correctes.

El que volem obtenir és una combinació entre les dues parts per aconseguir la generació de noves imatges, amb objectes totalment correctes i cohesionats i diferents de les imatges utilitzades per entrenar la màquina.

Així doncs l'objectiu d'aquest projecte és crear una màquina que ens permeti crear, a partir de dades correctes, moodboards, generar moodboards correctes sense la necessitat que una persona els hagi de crear, dissenyar un sistema automatitzat per a reemplaçar una tasca lenta i mecànica que realitzen a mà vàries persones.

A part de crear aquest sistema, també s'analitzarà la seva eficiència, respecte al sistema no automatitzat i si la generació d'aquests moodboards és correcta, així podrem establir si el sistema creat és prou eficaç, ràpid i precís a l'hora de generar, en aquest cas, moodboards.

Finalment, aquestes proves es realitzaran mitjançant la supervisió de persones expertes per a tal de millorar l'avaluació.

## Agraïments

Vull agrair a la empresa InteriorVista per donar-me l'oportunitat de realitzar aquest projecte i per tal de facilitar-me totes les eines i recursos necessaris per tal de poder desenvolupar i acabar el projecte.

També vull agrair al meu tutor a l'empresa, David Osuna, per proposar-me aquest projecte tant important i tant interessant per a realitzar i per ajudar-me en tots els dubtes que he tingut durant aquest temps.

I finalment vull agrair al meu tutor del treball, David Buchaca, per haver-me ajudat a organitzar el projecte i per haver-me ajudat a trobar eines i solucions per tal de resoldre els problemes que se'm presentaven al llarg del desenvolupament.

# Índex

<b>1</b>	<b>Introducció i objectius</b>	<b>1</b>
<b>2</b>	<b>Interfície gràfica per a moodboards</b>	<b>4</b>
2.1	Disseny dels moodboards . . . . .	4
2.2	Desenvolupament de la interfície gràfica i l'aplicació en Unity . . . .	8
2.2.1	Funcionalitats bàsiques . . . . .	8
2.2.2	Disseny i funcionament de l'aplicació . . . . .	10
2.2.3	Procés de generació d'un moodboard . . . . .	16
<b>3</b>	<b>Introducció al Machine Learning</b>	<b>18</b>
3.1	Història del Machine Learning . . . . .	18
3.2	Tipus d'algoritmes . . . . .	21
3.2.1	Aprenentatge supervisat . . . . .	21
3.2.2	Aprenentatge no supervisat . . . . .	22
3.3	Plantejament de l'algoritme . . . . .	22
<b>4</b>	<b>Boltzmann machine(BM) i Restricted Boltzmann machine (RBM)</b>	<b>25</b>
4.1	Boltzmann machine . . . . .	25
4.1.1	Entrenament Boltzmann machine . . . . .	26
4.2	Restricted Boltzmann machine . . . . .	28
4.2.1	Entrenament Restricted Boltzmann machine . . . . .	29
4.2.2	Avaluació d'una RBM . . . . .	32
4.2.3	Dificultats i solucions convencionals . . . . .	35
4.3	Algoritme seleccionat a implementar en codi . . . . .	36
<b>5</b>	<b>Desenvolupament de l'algoritme aplicat al problema</b>	<b>39</b>
5.1	Esquema del model . . . . .	39
5.2	Dades obtingudes de l'aplicació . . . . .	40
5.2.1	Format de les dades . . . . .	40
5.2.2	Dades obtingudes del projecte . . . . .	45
5.3	Dades d'entrenament . . . . .	45
5.3.1	Format de les dades . . . . .	46
5.3.2	Creació del codificador i dades obtingudes . . . . .	47

5.4	Creació i entrenament de la RBM . . . . .	49
5.4.1	Creació de la RBM a partir de les dades . . . . .	49
5.4.2	Entrenament de la RBM . . . . .	49
5.5	Dades de sortida . . . . .	51
5.5.1	Dades obtingudes de la RBM . . . . .	51
5.5.2	Creació de moodboards a partir de les dades . . . . .	53
<b>6</b>	<b>Resultats i conclusions i línies de continuació</b>	<b>55</b>
6.1	Qualitat de les graelles generades . . . . .	55
6.1.1	Possibles problemes . . . . .	57
6.2	Conclusions . . . . .	59
6.3	Treball futur: Com millorar el projecte . . . . .	61
	<b>Referències</b>	<b>62</b>

# 1 Introducció i objectius

Avui en dia, tot i que depenem de les màquines i dels programes per a realitzar tasques que poden ser pesades i lentes encara en queden moltes que es realitzen a mà. Moltes d'aquestes tasques en algun moment podrien ser automatitzades per tal que es puguin realitzar més ràpidament de forma menys pesada i probablement més precisa.

Aquest projecte vol aconseguir que una tasca que actualment costa temps, dedicació i concentració es transformi en una tasca menys tediosa i més ràpida de realitzar. L'objectiu del sistema que proposem és reduir el temps que un dissenyador triga a fer la tasca que volem optimitzar.

Aquesta tasca consisteix en la generació de moodboards. Un moodboard és una combinació d'imatges relacionades entre elles que juntes formen una sola imatge. Normalment els moodboards es fan servir per descriure combinacions de elements per a generar imatges de paisatges, d'aplicacions, decoració i creació d'habitacions etc...

El treball proposa facilitar la tasca de la generació de moodboards. Actualment per generar moodboards es necessita fer una cerca de totes les imatges que es desitgen col·locar a internet. Un cop triades les imatges s'han de col·locar de forma que formin un sol conjunt i anar-les editant per tal de generar una sola imatge. Aquest procés pot ser molt lent i tediós, ja que buscar totes les imatges necessàries costa temps. A més cal col·locar les imatges correctament i editar-les (retallant i enganxant des de programes de edició de imatges) per tal que totes s'ajuntin i formin una sola imatge.

Per a millorar la velocitat de generació de moodboards el treball es centra en dos aspectes:

- Crear una aplicació en unity, connectada a una base de dades, que permeti als dissenyadors buscar els ítems ràpidament en una interfície gràfica i arrastrar-los a una graella que determinara la moodboard. Aquesta aplicació ha de permetre buscar ítems amb característiques concretes com color de un objecte, que poden ser molt importants per a definir espais agradables.
- Usar un algorisme de machine learning per tal de generar noves moodboards a partir d'un conjunt de moodboards prèviament fet per dissenyadors. En cas que la qualitat de la generació no pugui competir amb la de un dissenyador aquest algorisme podria ser interessant per a generar un moodboard inicial des de la qual els dissenyadors poguessin començar a treballar.

## Hipòtesi

**L'objectiu del treball és validar si és possible incrementar la eficiència dels dissenyadors de moodboards amb una eina creada adhoc per a ells. En particular, voldríem validar si es possible crear un algorisme que faciliti la generació de moodboards usant Restricted Boltzmann Machines (RBM) per tal de donar una configuració inicial de ítems amb la que treballar.**

Per tal de poder crear aquest sistema utilitzarem tècniques de machine learning per tal de crear una màquina capaç d'aprendre i de generar a partir de les dades que ha après. D'aquesta forma podem estructurar el projecte en dues parts essencials. La primera part descriurà la generació de moodboards mitjançant una interfacial gràfica. La segona part descriurà la creació d'una màquina capaç de llegir els moodboards creats anteriorment i aprendre a crear-ne de nous.

Usarem les RBMs per a generar configuracions de moodboards de manera que un interiorista pugui començar a treballar des d'un moodboard "que tingui sentit", sense necessitat de començar des de zero. Poder fer això tindria un gran impacte en reduir el temps de cerca dels mobles i objectes necessaris que es necessiten per a cada zona d'una casa, obtenint moodboards d'una forma ràpida i eficient que tot i potser no tingui la qualitat per substituir el treball d'un interiorista pugui ser de prou qualitat com per a donar ja una bona feina feta.

Tot i que l'objectiu és crear un sistema automatitzat per a millorar la realització de la tasca de generació de moodboards també considerarem crucial analitzar com és de millor aquest sistema respecte de l'anterior. Per tal de saber-ho, al final de tot el projecte avaluarem el sistema mitjançant l'ajuda humana, en aquest cas, persones expertes que han generat molts de moodboards i que també han generat els correctes per entrenar la màquina i d'aquesta forma podrem validar les qualitats del sistema mitjançant diverses proves que s'explicaran més endavant.

## Estructura de la Memòria

En primer lloc ens centrarem com generem els moodboards, amb quines eines i alguns exemples per tal de crear-los. Aquests moodboards seran els que posteriorment serviran per entrenar la màquina. També explicarem el funcionament de les eines i les diferències amb el mètode “tradicional”.

Seguidament explicarem quin tipus de màquina farem servir, és a dir, el tipus d'algorisme d'aprenentatge automàtic aplicarem, els motius pels quals hem decidit escollir aquest algorisme i quines són les seves propietats bàsiques.

Després analitzarem com són les dades d'entrada, com a partir del moodboards generats obtenim, com transformem aquestes imatges per tal d'obtenir dades vàlides per a entrenar la màquina. També definirem com han de ser aquestes dades i com han de ser les dades que volem obtenir de sortida i com les transformarem en imatges de nou.

Per motius de confidencialitat la empresa no ha deixat publicar ni adjuntar el codi en aquesta memòria. De totes maneres, s'han inclòs parts de codi on es pot veure com s'han entrenat els models i els pre-processos necessaris per a generar els inputs.

Finalment, analitzarem els resultats obtinguts i els avaluarem mitjançant diferents proves que realitzarem per comprovar la seva eficiència i els resultats envers els moodboards creats prèviament per experts.

## Eines de programació usades

Per a poder realitzar aquest treball s'han usat les eines següents:

- Unity i C per a fer la interfície gràfica. La combinació dels dos softwares ens ha permès crear una eina fàcil de fer servir i intuïtiva que ens permet cobrir totes les característiques i funcionalitats demanades per aquest projecte.
- Python per a fer el model de generació de moodboards. S'ha escollit Python ja que compta amb una gran quantitat de biblioteques incorporades. Moltes de les biblioteques són per a intel·ligència artificial i aprenentatge automàtic. També ens permet tractar les dades de moltes maneres per tal d'obtenir els formats que es desitgen. Així doncs, és el llenguatge ideal per tal de desenvolupar la nostre màquina.



## 2 Interfície gràfica per a moodboards

Abans de començar a explicar les dades, el funcionament de l'algorisme i els resultats ens hem de centrar en la primera part del projecte i una de les més importants, la generació de dades.

Com hem comentat a la introducció les nostres dades són moodboards que defineixen diferents estances d'un habitatge d'un estil concret.

A continuació explicarem com són els moodboards, és a dir, com estan formats els moodboards actuals i com seran els nostres moodboards que generarem amb la nostra eina. També explicarem com està dissenyada la nostra eina, amb quines eines de programació s'ha realitzat i com es creen moodboards seguin l'estàndard que hem decidit.

### 2.1 Disseny dels moodboards

Com bé ja hem explicat un moodboard és una combinació d'imatges relacionades entre elles que juntes formen una sola imatge. Aquesta imatge que formen és una imatge ben generada, és a dir, que hi ha cohesió amb tots els elements de forma que sembli una fotografia mentre que es tracta una composició de diferents objectes editats manualment mitjançant eines d'edició d'imatge per formar una sola imatge coherent i compacta, la qual mostrarà una habitació d'una casa.

Realitzar moodboards amb aquesta tècnica té avantatges i inconvenients. Un dels problemes més grans que suposa és el temps. Generar un moodboard utilitzant aquest tipus de composició pot suposar una gran quantitat de temps, ja que en primer lloc necessitem buscar aquelles imatges que necessitem, és a dir, dintre d'un rang enorme el qual està format per milers de mobles i objectes escollir els adients segons l'estil i l'estança desitjada.

Un cop seleccionats els elements necessaris cal editar-los un per un per tal de poder combinar-los entre ells i que sembli una fotografia. Tot plegat suposa un esforç i una gran quantitat de temps per a generar un moodboard.

Per altra banda uns dels millors aspectes és la cohesió que s'aconsegueix i la impressió que provoca, ja que sembla una fotografia. Això el que ens permet és imaginar-nos molt millor com seria aquella habitació. Ens proporciona una visió molt més real que un conjunt d'imatges ajuntades aleatòriament.

La qüestió que ens plantegem és si la gran quantitat de temps invertida per a generar una sola imatge compensa el resultat obtingut. Per tal de respondre aquesta pregunta, hem consultat a experts que una de les seves principals tasques és la generació de moodboards.

Segons ells invertir una gran quantitat de temps per a generar una imatge compensa, ja que el resultat que obtenen és una reflexió de com seria veure en directe,

com fer una fotografia de l'habitació per tal de poder-la mostrar. També afirmen que a l'hora de mostrar habitacions a clients o empreses és una de les formes de captar més la seva atenció perquè en aquestes imatges es reflecteix la realitat del futur, és a dir, es reflecteix el que un dia podrà ser aquesta habitació en una casa qualsevol.

També han apuntat que en realitzar aquestes composicions els hi permeten crear diferents habitacions, diferents formes en ambients totalment diferents i poder afegir una història a darrere de cada composició que permeti a la persona que està veient aquella imatge entendre, endinsar-se dintre d'aquell ambient fictici fins a fer-lo real en la seva ment.

Així doncs, la generació de moodboards utilitzat aquesta tècnica genera molt bons resultats però té com a conseqüència dedicar-hi molt de temps per tal d'obtenir un bon resultat i que captivi aquelles persones a les quals va dirigit.

A partir d'aquest punt el que ens hem de plantejar si utilitzar l'esquema de la tècnica anterior és adient pel nostre projecte, és a dir, si ens podem permetre generar una gran quantitat d'imatges dedicant-hi molt de temps per després utilitzar-les per entrenar una màquina.

La resposta és simple, aquest tipus de tècnica no serveix per a aquest cas, ja que necessitem generar, amb un període curt de temps, una gran quantitat de moodboards. Per tal de solucionar-ho, s'ha de plantejar una solució que ens permeti organitzar en un moodboard els elements d'una forma ràpida i que més o menys formin una imatge cohesionada.

Partint d'aquesta premissa es van agafar diferents estudis realitzats sobre moodboards anteriorment generats de cada estança i es van treure una sèrie d'especificacions que cada moodboard hauria de complir, així podríem definir un model, una plantilla per a generar moodboards d'una forma més mecànica i més ràpida.

Així doncs una de les especificacions més importants que es va arribar és que cada estança que es volia generar tenia una sèrie d'elements fixes, és a dir, una sèrie de mobles i objectes que sempre hi haurien de ser. D'aquesta forma es va establir un nombre d'objectes per a cada estança i quins havien de ser. Això solucionava una part del problema que era saber quants objectes podíem col·locar a cada moodboard. L'altre problema que va sorgir era com combinar imatges per tal d'aconseguir una imatge cohesionada semblant a les imatges que s'obtenien amb l'altra tècnica. En aquest cas el rang d'objectes disponible era bastant gran, es tractava d'un conjunt d'uns 3000 objectes, i les imatges per a cada objecte eren *thumbs*, imatges amb un fons blanc i l'objecte en petit centrat al mig del fons blanc. La solució de treure el fons blanc editant la imatge no era viable, ja que es tardaria molt i no era l'objectiu que volíem aconseguir.

Partint de tot això, es va arribar a trobar una solució que complia amb les especificacions i més o menys s'obtenia un resultat coherent. Partint de les solucions anteriors es va establir una graella per a cada estança. Aquesta graella contindria a cada casella un tipus d'objecte, la qual més endavant contindria la imatge, el

*thumb* de l'objecte que fos d'aquell tipus. D'aquesta forma s'aconsegueix una certa cohesió, ja que totes les imatges en tenir el mateix format es formaria una imatge de fons blanc amb petites imatges de cada objecte. Tot i això, sempre hi ha la condició que a una cel·la no hi hagi cap element i quedi de color blanc.

Aquesta solució serviria per a totes les estances disponibles d'un habitatge però si ens centrem amb el projecte, només analitzarem una estança, el dormitori per tant en aquest cas tindrem una sola graella d'imatges que corresponent als objectes que es troben en l'estança. Així doncs es va establir que, en aquest cas, la graella del dormitori tindria 48 elements els quals correspondrien als diferents mobles i objectes que es poden col·locar.

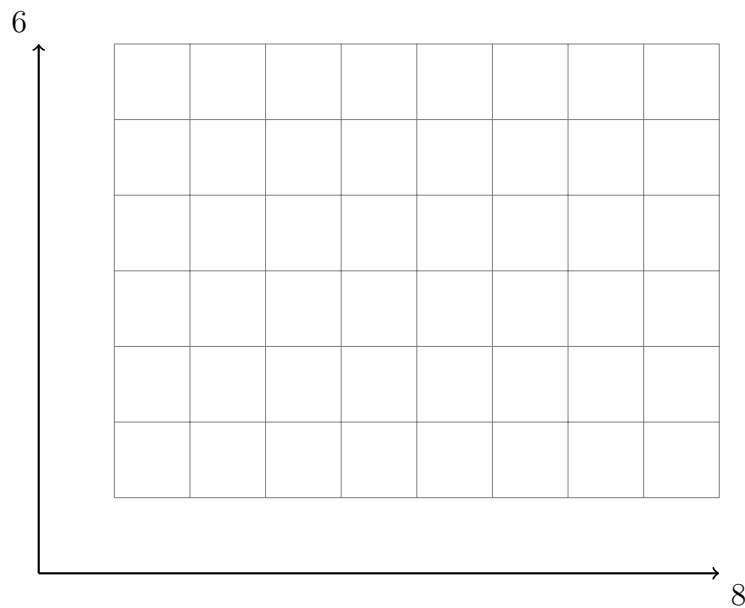


Figura 1: Exemple de la graella de l'estança dormitori.

Cada casella d'aquesta graella tindrà internament, és a dir, d'una forma que no es veurà visualment uns valors que indicaran quins objectes poden anar aquella casella. D'altra banda cada casella també contindrà una o més d'una paraula que descriurà els valors anteriors, d'aquesta forma la persona que hagi d'emplenar la graella amb objectes ha de saber quins poden i quins no. A continuació es mostra la graella final que s'utilitzarà per a la generació de moodboards.

Suelo	Pared A	Pared B	Resto Paredes	Pilar o Columna	Murete	Cornisa	Molduras Frisos
Arrimadero	Zócalos	Cantonera	Escaleras	Techo	Viga o Estructura Compuesta Vigas	Molduras	Rosetón
Puerta	Ventana o Balconera	Otros	Separadores Espaciales	Calefacción	Climatización Ventilador Techo	Chimenea	Interruptores Enchufes
Iluminación Técnica	Cama	Estantería encima cama	Mesita Noche	Lámpara Techo	Armario	Banco/Baúl	Colgadores
Cómoda	Espejode Pie	Galán	Butaca	Lámpara de pie	Mesita Auxiliar	Tocador	Espejo
Alfombra	Cortinas	Textil Cama Principal	Textil Cama Secundario	Textil Cama Secundario 1	Textil Cama Secundario 2	Escritorio	Estantes de Pared

Figura 2: Graella final de l'estança dormitori.

A partir d'aquest punt ja tenim una solució que ens permet generar moodboards de forma ràpida encara que el resultat no sigui el mateix que utilitzant la primera tècnica descrita. Tot i això, en aquest cas preferim velocitat a l'hora de generar que més cohesió, ja que tampoc tenim forma que un cop la màquina generi graelles pugui editar aquestes imatges per tal que s'assemblin al model de la primera tècnica. A partir d'ara necessitem una forma de poder col·locar els objectes disponibles a la graella d'una forma senzilla i eficient.

## 2.2 Desenvolupament de la interfície gràfica i l'aplicació en Unity

Per tal de poder aplicar el mètode anteriorment explicat necessitem alguna eina que ens permeti accedir a tots els objectes disponibles d'una forma fàcil per tal de poder crear moodboards seguint la graella prèviament definida. Així doncs necessitem crear una eina per realitzar aquesta tasca.

Aquesta eina que volem desenvolupar ha de ser fàcil de fer servir, intuïtiva i ràpida. Partint d'això s'ha de decidir quines tecnologies utilitzar per desenvolupar-la. En aquest cas, com que ja s'havien creat altres eines per altres coses es va decidir aplicar les mateixes tecnologies que en aquest cas són utilitzar Unity i el llenguatge de programació C sharp.

Aquestes dues tecnologies combinades ens permetran desenvolupar una eina capaç de generar moodboards d'una forma ràpida i eficient.

En primer lloc tenim la part de Unity que ens permetrà crear tot el disseny de la interfície i les interaccions necessàries perquè l'usuari pugui col·locar els objectes i així generar moodboards. També ens permetrà visualitzar tots els objectes i tota la informació que necessitem per a la generació.

I en segon lloc la part de C Sharp que controlarà tot la lògica que hi ha darrere de la interfície. Però el paper més important és la connexió amb la base de dades. Els objectes que tenim disponibles, la seva informació, la graella creada i tots els moodboards generats es guarden en una base de dades que es troba en un servidor. Així doncs per tal de poder accedir a aquesta informació, s'han realitzat crides constants per obtenir-la i això se n'encarrega la part lògica de l'aplicació.

Pensant en un esquema model-vista-controlador (MVC), la part de model i vista correspondrien a Unity mentre que tota la part del controlador seria gestionada per fitxers, scripts, amb C Sharp els quals s'han de comunicar amb la base de dades per tal de proporcionar la informació necessària per a mostrar tots els objectes, la graella i altres funcions que ja explicarem.

### 2.2.1 Funcionalitats bàsiques

Com hem comentat aquesta eina ha de tenir unes funcionalitats bàsiques i essencials per a obtenir un bon resultat. Aquestes funcionalitats són les següents:

- Categoritzar graella predefinida. Com hem dit anteriorment, la graella la qual estableix quin objecte pot anar a cada casella, està organitzada de forma que cada una d'elles conté una sèrie d'identificadors els quals cada un és una subcategoria, un tipus d'objecte de la base de dades. Aquests identificadors no es veuen a la graella, ja que no tenen importància visualment però el que sí que es veu és un resum d'aquestes categories que especifica de forma més general la casella.

Aquests noms i aquests valors s'han de poder modificar per si en algun moment s'ha classificat una casella incorrectament. Així doncs, l'eina ens ha de permetre modificar cada una de les caselles, tant la part visual com els valors que no es mostren que en permeten filtrar objectes.

- Carregar objectes. Com hem mencionat repetidament necessitem objectes per a col·locar. Aquests objectes es troben a una base de dades i el que l'eina ha de fer és obtenir la informació d'aquests i mostrar-la en forma d'imatges per tal que l'usuari pugui col·locar-los correctament a la graella. Així doncs ha d'implementar un sistema de crides a servidor per obtenir les dades i transformar-les per tal que es mostrin correctament i s'emmagatzemi d'alguna forma la informació que posteriorment ens serà útil.
- Filtrar objectes. El rang disponible d'objectes és molt ampli i el que no volem és que l'usuari es passi molt de temps buscant un objecte per a col·locar-lo a una casella. Per tant, el que es necessita és implementar un sistema de filtratge ràpid i fàcil de fer servir.  
Per tal d'implementar aquest sistema utilitzarem els valors que amaguen les caselles que són els que ens indiquen aquells objectes que poden anar en ella. Així doncs cada cop que es vulgui col·locar un objecte a una casella només es mostraran aquells objectes que la base de dades retornarà aplicant el filtre amb els valors de la casella. D'aquesta forma aconseguim que l'usuari no s'hagi de trencar al cap buscant objectes i sempre col·loqui objectes vàlids a cada casella.  
A part també s'han de poder filtrar aquests objectes, els que ja han estat filtrats, segons cada valor de la casella, és a dir, ha d'haver-hi l'opció de mostrar tots els objectes disponibles de la casella o només els d'un tipus determinat de la casella, només utilitzant un valor de la casella.
- Col·locar objectes. L'aplicació ha de permetre a l'usuari poder col·locar objectes a la graella. Així doncs, la mecànica és que pugui arrossegar la imatge de l'objecte a la graella. També ha de permetre que un cop col·locat l'objecte aquest es pugui treure i/o eliminar.
- Guardar graella. Com que el principal objectiu de l'aplicació és generar moodboards per després entrenar una màquina amb els moodboards generats, l'aplicació ens ha de permetre guardar d'alguna forma el moodboard generat amb tota la distribució d'objectes creada per l'usuari. El format d'aquestes dades que s'exportaran ha de ser l'adient per tal que la màquina pugui llegir-les.
- Crear PDF. Tot i que només necessitem exportar les dades per la màquina també s'ha demanat una forma de visualitzar els moodboards generats i la

informació que contenen. Així doncs, l'aplicació ha de permetre generar d'algun forma un fitxer on figuri tota la informació. La solució és generar un PDF. Aquest fitxer, en primer lloc contindrà una captura de la graella i els seus elements i a continuació a cada pàgina contindrà l'objecte amb tota la seva informació, així obtenim un fitxer fàcil de compartir i el qual conté tota la informació del moodboard generat.

## 2.2.2 Disseny i funcionament de l'aplicació

A continuació explicarem el disseny de l'aplicació, és a dir, com és la interfície gràfica, per a què serveix cada part que se'ns mostra. Per explicar-ho millor partirem d'una imatge de l'eina, ja que ens servirà per anar assenyalant millor les parts i indicant que fa cada part. Més endavant explicarem quin procés s'ha de seguir per a crear un moodboard. A continuació es mostra la imatge de l'eina.

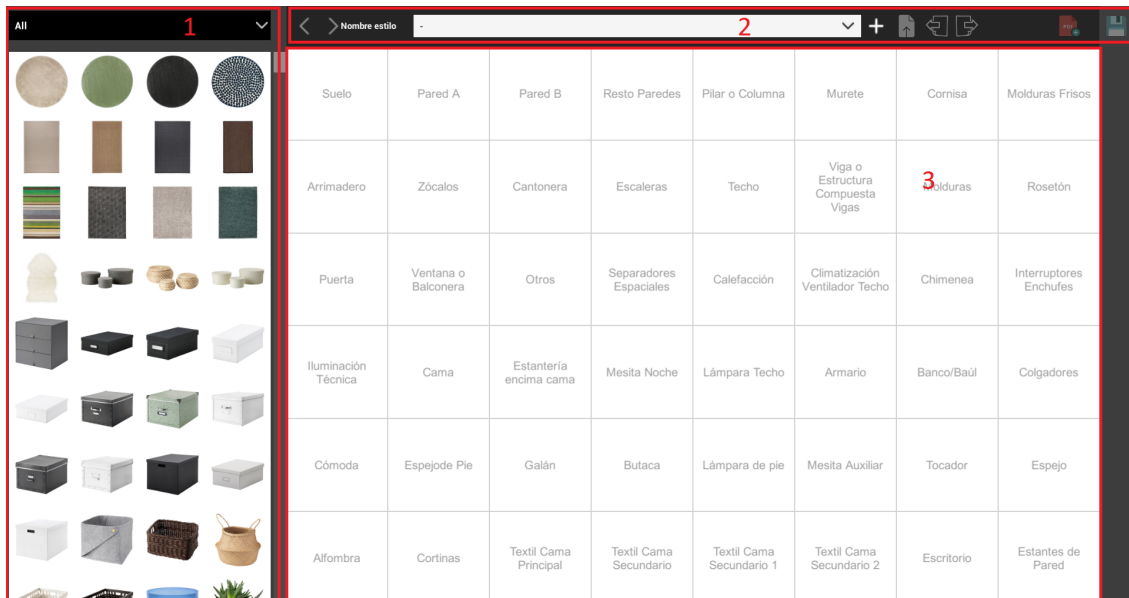


Figura 3: Captura de l'eina.

Podem observar que es tracta d'una interfície senzilla i ben organitzada. A l'hora de dissenyar l'eina es van tenir en compte totes les normes de disseny i d'usabilitat per tal que la distribució de la informació i dels botons fos la més adient per tal que l'usuari pogués treballar d'una forma ràpida i mecànica.

Com podem observar l'eina es divideix en tres parts essencials els quals cadascuna correspon alguna de les necessitats més importants que s'han mencionat anteriorment. A continuació analitzarem les 3 zones per separat explicant el seu objectiu i el funcionament dintre de l'eina.

Primer de tot començarem per la zona 3, la graella. Aquesta graella és la mateixa que s'ha mostrat prèviament i s'ha explicat com estava formada. Podem observar que aquesta graella és prou gran perquè en col·locar un objecte aquest no sigui ni molt petit ni molt gros sinó la mida justa per tal que es pugui veure de quin objecte es tracte. Cada casella de la graella és clicable, és a dir, podem fer clic sobre la casella i aquesta quedarà ressaltada de color verd. El que ens permet és que en seleccionar una casella els objectes es filtrin segons les categories de la casella.

Una altra funcionalitat de les caselles és que si en comptes de fer clic amb el botó esquerre del ratolí el realitzem amb el botó dret, eliminarem l'objecte que es troba a la casella. Aquest funcionament permet a l'usuari una ràpida manipulació de les graelles per tal de poder trobar els objectes adients per a cada una i la facilitat d'eliminar i/o modificar sense la necessitat de realitzar més moviments dels necessaris.

Tal com hem mencionat abans la graella es pot modificar, es poden canviar les categories. Per tal de modificar-la necessitem entrar en el mode edició, en aquest mode, en fer clic amb el botó esquerre ens apareixerà una finestra la qual haurem d'indicar el nom que volem que es mostri i escollir les subcategories que desitgem. Per a escollir una subcategoria primer s'ha de seleccionar el tipus, després la categoria i després la subcategoria.

Per entrar i sortir del mode edició l'usuari només ha de prémer la tecla "F1". Un cop acabada l'edició de les graelles, es guardarà la configuració i es mantindrà fins que no torni a ser modificada de nou per l'usuari. A continuació es mostra una imatge de la finestra que apareix quan ens trobem en el mode edició.



Name: Suelo

Type: ▾

Category: ▾

Subcategory: ▾ +

**Elementos Estructurales/Pavimento/Compuesto**

**Elementos Estructurales/Pavimento/Continuo**

Eliminar Selección      Cancelar      Guardar

Figura 4: Captura de la finestra del mode edició.



La graella es pot emplenar del tot o deixar caselles buides. No hi ha cap restricció sobre els elements mínims necessaris per a generar un moodboard.

Un cop ja hem explicat l'element principal de l'eina a continuació ens centrarem amb la zona 1, la zona dels objectes. Com hem dit, necessitem visualitzar els diferents objectes que es troben a la base de dades, així doncs necessitem una zona en la qual podem veure i filtrar els diferents objectes.

Només d'iniciar l'aplicació els objectes que es mostraran són tots els objectes disponibles que tenim, ja que no s'ha seleccionat cap cel·la en concret. El rang d'objectes que tenim és molt gran i cada objecte conté una imatge que l'aplicació ha de descarregar. Això ens presenta un problema important que haurem de gestionar d'una forma especial.

El problema que se'ns presenta és que no podem carregar tots els objectes amb les seves corresponents imatges, ja que aquesta càrrega trigaria molt i podria fer lent la tasca de generació de moodboards.

Per tal de solucionar aquest problema crearem un sistema de càrrega semblant als buscadors de les pàgines web, les imatges s'aniran carregant a mesura que es mogui la barra de desplaçament o s'utilitzi la rodeta del ratolí per baixar la barra. Aquest sistema ens permet una càrrega ràpida i fluida de les imatges sense la necessitat de carregar totes les imatges. Així doncs, en iniciar l'aplicació es mostraran un nombre finit d'imatges, aquest número dependrà de la mida de la finestra, ja que com més gran més imatges es mostraran al principi, i a mesura que es vagi desplaçant cap avall s'aniran carregant les següents imatges.

També s'ha implementat que un cop la imatge s'ha descarregat no es torni a descarregar, es vol imitar el sistema de memòria caché però en aquest cas les imatges es guarden internament en l'aplicació el que ens permet que un cop carregada una imatge ja no es tornarà a descarregar, conseqüentment, si s'han carregat tots els objectes amb les seves imatges, ja no es descarregaran més imatges i el filtratge d'objectes serà més ràpid. Aquesta informació desapareixerà un cop l'aplicació es tanqui.

El filtratge d'objectes es realitzarà quan se seleccioni una cel·la. Les diferents opcions de filtratge es mostraran en el desplegable que es troba a la part superior. En tots els casos hi haurà una opció que serà el filtratge de totes les categories, és a dir, de tots els objectes que poden anar a la cel·la i les altres opcions seran per filtrar per categoria, per escurçar el rang d'objectes a col·locar. El sistema de càrrega d'imatges del filtratge és el mateix que s'ha explicat anteriorment, ja que si ja tenim imatges descarregades només les hem de mostrar i en cas contrari, les descarreguem i les guardem per posteriors usos.

Tot el filtratge d'objectes s'obté mitjançant crides a la base de dades aplicant els filtres que té la cel·la. Un altre sistema seria, filtrar a partir de la llista d'objectes descarregats inicialment, però això pot suposar que en algun moment s'afegeixi o es modifiqui un objecte mentre s'està utilitzant l'aplicació i aquest no apareixeria

en el filtratge per això utilitzem les crides a base de dades per obtenir la informació actualitzada. A continuació es mostra un exemple de com funciona el filtre un cop se selecciona una cel·la. Podem veure com la primera opció és la general, la que engloba les dues categories i després podem filtrar per cada una d'elles.

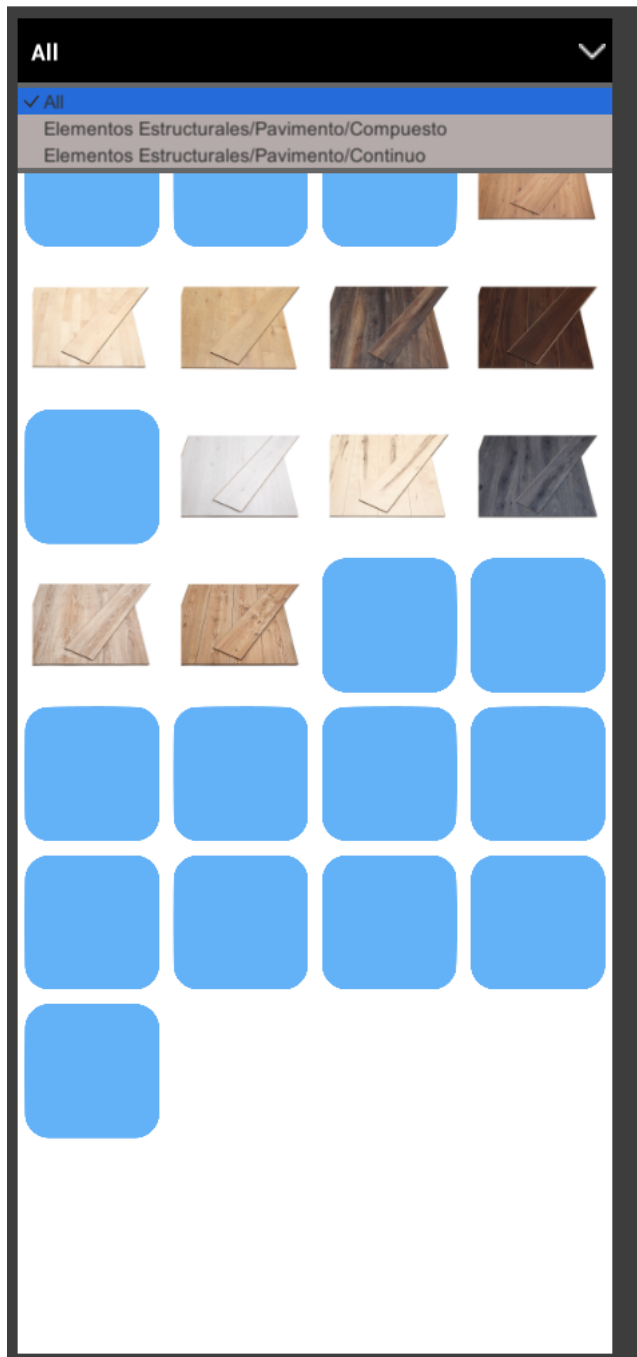


Figura 5: Captura del filtratge després de seleccionar una cel·la.

Un cop ja explicades dues de les parts de l'eina, finalment explicarem la part superior la qual es compon dels diferents botons de l'aplicació per a realitzar diferents

accions. A continuació es mostra una imatge de la barra superior de botons. A partir d'aquestes imatges explicarem la funcionalitat de cada botó a dins l'eina i les diferents parts que la formen.



Figura 6: Barra superior de l'aplicació.

Seguint un ordre d'esquerra a dreta tenim en primer lloc dos botons que són els botons de desfer i refer. L'aplicació té implementat un sistema de gestió d'accions, és a dir, cada cop que es realitza una acció aquesta queda registrada a dins l'aplicació. Això el que ens permet és tenir un control sobre elles i en qualsevol moment saber quines accions s'han realitzat.

A partir d'aquest sistema de gestió hem implementat un sistema de desfer i refer que desfà o torna a fer les accions que s'han realitzat. Això permet que si en un punt ens hem equivocat i volem tornar endarrere no tinguem problema o si s'ha guardat però ens hem equivocat puguem tornar a un punt anterior per seguir endavant.

Aquest sistema guarda totes les accions que es realitza l'usuari amb el ratolí. Aquestes dues icones s'aniran il·luminant a mesura que sigui possible realitzar les accions, és a dir, si no es pot retrocedir o no es pot refer, no es podrà interactuar amb el botó, així s'evita que es realitzin accions no permeses.

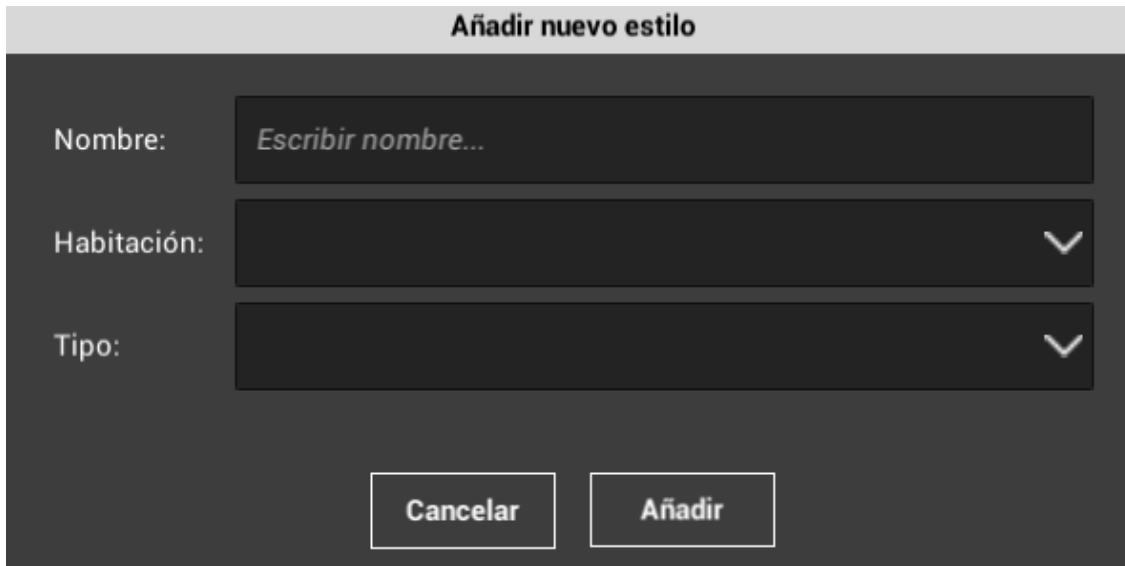
Cada cop que es reinicia l'aplicació és reseteja el sistema i es netegen les accions, quan es guarda un moodboard definitiu també es netegen tot i que es pot seguir realitzant accions que seran afegides al sistema com a noves accions, és a dir, com si es tornés a crear el sistema de gestió. Això permet indicar a l'usuari que ha guardat el que ha realitzat i que a partir d'aquest punt tot el que realitzi serà un nou moodboard.

Després dels botons de desfer i refer tenim el nom del moodboard. Aquí és on ens apareixerà el nom del moodboard que estem generant. En el desplegable també ens apareixerà el nom dels moodboards anteriors que hem creat. Seleccionar un nom d'un moodboard que ja hem creat ens permet crear un nou moodboard amb el mateix o carregar-ne un de nou per a visualitzar-lo.

Per tal de crear un nou moodboard hem de prémer el botó amb el símbol més. En prémer aquest botó ens apareixerà una nova finestra en el qual ens demanarà que entrem unes dades. En primer lloc en demanarà el nom de l'estil, que potser qualsevol nom que l'usuari desitgi. Després utilitzant un desplegable l'usuari ha de seleccionar a quina habitació correspon l'estil i finalment haurà de seleccionar el tipus que correspon.

El tipus es refereix a la distribució d'attrezzo, és a dir, la distribució d'objectes

que es podrien posar a cada moble, a sobre, a sota, els complements per a cada moble que es podrien col·locar per acabar de completar el moodboard. Hi ha sis tipus respecte si estan ordenats o poc i amb quina quantitat. Quan s'afegeix un nou moodboard aquest s'afegirà automàticament al desplegable. A continuació es mostra la pantalla per afegir un nou moodboard.



The image shows a dark-themed dialog box titled "Añadir nuevo estilo". It contains three input fields: "Nombre:" with a text input field containing the placeholder "Escribir nombre..."; "Habitación:" with a dropdown menu; and "Tipo:" with a dropdown menu. At the bottom, there are two buttons: "Cancelar" and "Añadir".

Figura 7: Barra superior de l'aplicació.

A continuació segueix el botó per a carregar moodboards ja creats. Això ens permet visualitzar moodboards creats prèviament. Per tal de poder carregar un moodboard correctament s'ha de seleccionar un moodboard. Per fer-ho només cal obrir el desplegable i seleccionar el nom del moodboard desitjat. Si no se selecciona cap nom, la icona no estarà activa, ja que seleccionar un nom és un requisit indispensable per tal de fer la càrrega correctament.

Un cop es premi el botó apareixeran els objectes a la graella segons la distribució del moodboard carregat. L'aparició dels objectes pot ser una mica lenta, ja que per a cada casella ha de buscar l'objecte que hi ha d'anar i si aquest no té imatge descarregar-la. Aquest procés pot ser lent, però si tenim en compte que prèviament s'han carregat totes les imatges de tots els objectes, la col·locació d'aquest ha de ser ràpida.

Com hem dit anteriorment, un moodboard pot contenir diferents distribucions d'objectes, així doncs, en carregar aquest moodboard també hem de poder visualitzar totes les distribucions disponibles. Els dos botons del costat del botó de carregar moodboards ens permetran mostrar les diferents distribucions que té el moodboard. Només de carregar el moodboard es mostrarà la primera distribució de totes. Si només es disposa d'una distribució, els dos botons es trobaran inactius mentre que si hi ha més d'una distribució els botons s'aniran activant per permetre l'avanç

o el retrocés de les distribucions. En qualsevol dels casos, les distribucions si per algun motiu està malament o s'han de modificar és poder canviar els objectes que les formen. En cas de modificació, la distribució es guardarà segons el nom escollit, si es dóna el cas que és el mateix nom, s'afegirà a la distribució corresponent i aquesta podrà ser carregada i visualitzada més endavant.

Seguidament ens trobem amb el botó de generació de PDF. Aquest botó ens permetrà generar un PDF del moodboard actual en el qual apareixerà una captura de la graella amb els elements, amb el nom del moodboard i a continuació a cada una de les pàgines del fitxer, cada objecte amb la seva informació corresponent.

En fer clic sobre el botó ens apareixerà una finestra en la qual ens indicarà a on volem guardar el PDF que es generarà. Un cop seleccionada la ruta només hem de prémer el botó de guardar i a continuació es generarà el PDF. Aquest procés pot tardar una miqueta, ja que generar un fitxer amb molta informació pot ser lent, d'aquesta forma, per indicar que el fitxer s'ha generat correctament, la icona es desactivarà, indicant que s'ha creat el fitxer correctament i que el PDF del moodboard actual ja s'ha generat. En cas de modificació del moodboard, es tornarà a activar la icona per a generar un nou fitxer del moodboard.

Finalment trobem el botó de guardar. Aquest botó ens permet guardar el moodboard generat de forma que s'exporta en un format que després serà tractat per utilitzar-lo amb la màquina. Per tal de poder guardar un moodboard és necessari que aquest tingui un nom si no la icona de guardar no es trobarà activa.

Si en guardar, no existeix cap moodboard amb el nom escollit, es crearà un nou fitxer en el qual s'emmagatzemaran els moodboards amb el mateix nom. En cas contrari, que ja existeixin moodboards amb el mateix nom, s'afegirà la distribució d'objectes al fitxer. D'aquesta forma, per un sol moodboard podem tenir diferents distribucions.

S'ha desenvolupat d'aquesta forma per tal d'assegurar que el format en el qual cada moodboard és exportat és fàcilment manipulable per transformar-lo per un posterior tractament per entrenar la màquina.

### 2.2.3 Procés de generació d'un moodboard

En aquest punt explicarem els passos a seguir per crear un moodboard correctament, per a generar un moodboard pas a pas utilitzant tota la informació prèviament explicada. A continuació s'enumerarà cada pas i en acabat es mostrarà un diagrama de seqüència per tal d'entendre-ho millor.

1. **Obrir correctament l'aplicació i comprovar que es carreguen tots els objectes i la graella.** En cas contrari pot ser culpa de la connexió a internet o la connexió a base de dades.
2. **Categoritzar la graella.** En cas que la graella estigui buida categoritzar-la per tal de fer el filtratge d'objectes correctament.

3. **Escollir un nom o afegir-ne un de nou.** Escollir un nom pel moodboard per tal de poder guardar i generar el PDF en qualsevol moment.
4. **Emplenar graella.** Emplenar la graella amb els objectes necessaris fins a obtenir un moodboard correcte. Utilitzar les funcions de filtratge per tal de poder col·locar els objectes amb més precisió i més eficaçment.
5. **Guardar moodboard.** Un cop s'hagi acabat d'emplenar el moodboard guardar el moodboard correctament per tal de no perdre els canvis realitzats.
6. **Generar PDF.** Un cop guardat el moodboard guardar el PDF per tenir una referència i una visualització ràpida del moodboard generat.

A continuació es mostra un esquema de seqüència dels passos a realitzar de completar els passos anteriors.

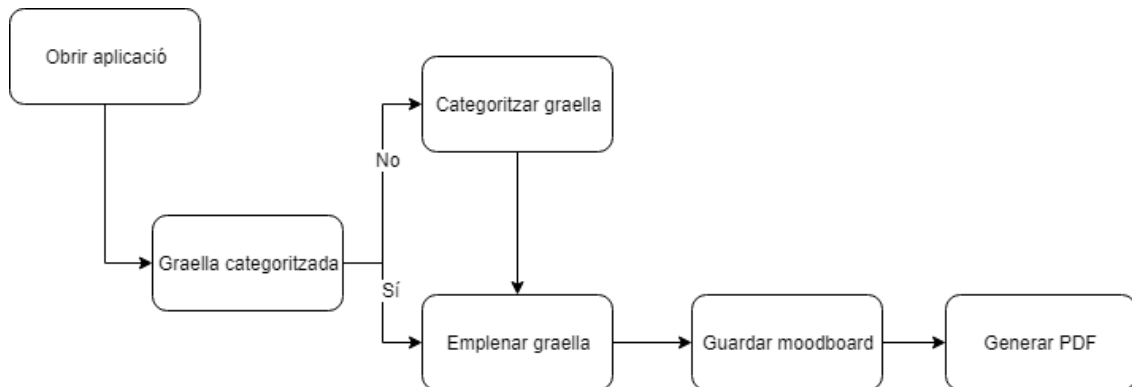


Figura 8: Digrama de seqüència per a generar un moodboard.

## 3 Introducció al Machine Learning

La part més important de projecte no és el desenvolupament de l'eina i el seu funcionament, sinó que és la part del algorisme. L'algorisme és el que ens permetrà generar automàticament moodboards un cop aquest s'hagi entrenat i és el que ens permetrà, que la generació de moodboards sigui ràpida i automàtica.

Així doncs, a continuació explicarem en què es basa l'algorisme que realitzarem, explicant una mica d'introducció i història per posar en context. També explicarem els diferents tipus i quin d'ells s'ha escollit i com s'ha desenvolupat.

### 3.1 Història del Machine Learning

Parlar avui d'Intel·ligència Artificial no és una novetat. Ni per a la gent ni per a les empreses o governs.

Si bé sembla una qüestió d'allò més naturalitzada per aquests dies, fins fa no molts anys enrere, parlar d'intel·ligència artificial era un assumpte distant, difícil d'abordar, reservat només per a entesos.

Avui en dia, es parla i s'aplica per qüestions tan bàsiques com l'ús del homebanking, veure pel·lícules en streaming o escoltar música.

Cada vegada més, es llegeix i es comenta sobre intel·ligència artificial en mitjans massius de comunicació, amb tota facilitat. Fins i tot, comença a guanyar un espai dintre del pressupost propi de les empreses al país: la recerca per la reducció dels costos i l'augment de la productivitat vénen impulsant l'avanç en l'ús de solucions que involucren robots capaços de conversar amb els clients i de sistemes que analitzen milers de dades en pocs segons.

Dins de l'ampli espectre que abasta la intel·ligència artificial, es troba Machine Learning. Però que és el Machine Learnig?

El machine learning, conegut com aprenentatge automàtic, va néixer com una idea ambiciosa de la IA en la dècada dels 60. Per ser més exactes, va ser una subdisciplina de la IA, producte de les ciències de la computació i les neurociències.

El que aquesta branca pretenia estudiar era el reconeixement de patrons (en els processos d'enginyeria, matemàtiques, computació, etc.) i l'aprenentatge per part de les computadores. En les albers de la IA, els investigadors estaven àvids per trobar una forma en la qual els ordinadors poguessin aprendre únicament des de dades.

Va succeir amb el pas dels anys que el machine learning començar a enfocar-se en diferents assumptes, com ara el raonament probabilístic, investigació basada en l'estadística, recuperació d'informació, i va continuar aprofundint cada vegada més en el reconeixement de patrons (tots aquests assumptes aplicats a processos d'enginyeria, matemàtiques, computació i altres camps relacionats amb objectes físics o abstractes).

Això va ocasionar que en els 90 es separés de la IA per convertir-se en una disciplina per si sola, encara que molts puristes encara la consideren com a part de la IA. Ara, el principal objectiu del machine learning és abordar i resoldre problemes pràctics en on s'apliqui qualsevol de les disciplines numèriques abans esmentades. Com vam establir prèviament, és un camp de les ciències de la computació que, d'acord a Arthur Samuel el 1959, li dona a les computadores l'habilitat d'aprendre sense ser explícitament programades.

Si aquesta definició va resultar molt trivial, posem-d'aquesta manera: és la idea que hi ha algoritmes que poden donar-te troballes o conclusions rellevants obtingudes d'un conjunt de dades, sense que l'ésser humà hagi d'escriure instruccions o codis per això.

El propòsit del machine learning és que les persones i les màquines treballin de la mà, aquestes màquines han de ser capaços d'aprendre com un humà ho faria. Precisament això és el que fan els algoritmes, permeten que les màquines executin tasques, tant generals com específiques. Si bé al principi les seves funcions eren bàsiques i es limitaven a filtrar emails, avui dia pot fer coses tan complexes com prediccions de trànsit en interseccions molt transitades, detectar càncer, mapejar llocs per generar projectes de construcció en temps real, i fins i tot, definir la compatibilitat entre dues persones.

El principal objectiu de tot aprenent (learner) és desenvolupar la capacitat de generalitzar i associar. Quan traduïm això a una màquina o ordinador, significa que aquestes haurien de poder exercir-se amb precisió i exactitud, tant en tasques familiars, com en activitats noves o imprevistes.

I com és possible això? Fent que repliquin les facultats cognitives de l'ésser humà, formant models que "generalitzin la informació que se'ls presenta per realitzar les seves prediccions. I l'ingredient clau en tota aquesta qüestió són les dades.

En realitat, l'origen i el format de les dades no és tan rellevant, ja que el machine learning és capaç d'assimilar una àmplia gamma d'aquests, el que es coneix com big data, però aquest no els percep com a dades, sinó com una enorme llista d'exemples pràctics.

Podríem dir que les seves algoritmes es divideixen principalment en tres grans categories: supervised learning (aprenentatge supervisat), unsupervised learning (aprenentatge no supervisat) i Reinforcement learning (aprenentatge per reforç). A continuació, detallarem les diferències entre aquestes.

Una vegada que comproves el fàcil i pràctic que resulta d'aplicar les tècniques de machine learning a problemes que creies serien impossibles, és quan comences a creure que podria resoldre pràcticament qualsevol problema, sempre i quan hi hagi suficients dades ja que la falta de dades pot provocar problemes al desenvolupament.

Per al consumidor modern, el machine learning és un facilitador clau de moltes de



les seves tasques quotidianes. Des serveis de traducció, a prediccions climàtiques, fins endevinar el que els usuaris volen amb base a les seves activitats recents; les prestacions que ofereix són incomparables.

Pel que fa als negocis, moltes companyies han començat a incorporar aquesta tecnologia als seus sistemes operatius, amb grans expectatives de millorar i automatitzar els seus processos.

Atès que el machine learning és un sistema basat en el processament i anàlisi de dades que són traduïts a troballes, es pot aplicar a qualsevol camp que compti amb bases de dades prou grans. De moment, alguns dels seus usos més populars i desenvolupats són:

- Classificació de seqüències de DNA
- Prediccions econòmiques i fluctuacions en el mercat borsari
- Mapatges i modelats 3D
- Detecció de fraus
- Diagnòstics mèdics
- Cercadors a Internet
- Sistemes de reconeixement de veu
- Optimització i implementació de campanyes digitals publicitàries

Tot i que ja hem vist el que la IA és capaç d'aportar a les nostres activitats del dia a dia, com podria això beneficiar el món dels negocis?

Bé, ja que les converses i comentaris d'una infinitat de consumidors digitals li ofereixen a aquest tipus de tecnologies una quantitat d'informació aclaparadora, aquestes contínuament obtenen coneixements nous i detecten tendències més ràpid del que qualsevol humà podria fer-ho.

Si bé és cert que aquesta enorme quantitat de dades la tornarà molt més eficient, requerirà necessàriament de molt talent humà per perfeccionar-se, ja que finalment els ordinadors no tenen un domini tan elevat del llenguatge aplicat al raonament. O el que és, no són precisament hàbils per determinar contextos.

El que significa que perquè el machine learning es desenvolupi en aquestes àrees, els experts en cada camp de treball hauran de trobar el temps per entrenar a les màquines i anar incorporant paulatinament a cada un dels processos que desitgin afinar.

Finalment s'estima que aquesta -com molts altres derivats de la IA- transformarà per complet el món com el coneixem. Com podem observar això només reflecteix una petita introducció al intricat món del machine learning. En una època on

emergeixen tecnologies innovadores cada vegada que parpellegem, és fàcil perdre en l'allau d'informació i nous conceptes.

## 3.2 Tipus d'algoritmes

Dintre del conjunt del Machine Learning poder distingir entre diferents tipus que ens definiran diferents tipus d'algoritmes a aplicar durant el desenvolupament d'una màquina. Els principals tipus que tractarem per aquest projecte són els següents:

- Aprenentatge supervisat
- Aprenentatge no supervisat

A continuació es descriuen breument el aprenentatge supervisat i el no supervisat. Seguidament mostra una imatge resum dels diferents tipus i un exemple de cada un.

### 3.2.1 Aprenentatge supervisat

En l'aprenentatge supervisat, els algorismes treballen amb dades etiquetades (labeled data), intentant trobar una funció que, donades les variables d'entrada (input data), els assigni l'etiqueta de sortida adequada. L'algorisme s'entrena amb un històric de dades i així aprèn a assignar l'etiqueta de sortida adequada a un nou valor, és a dir, prediu el valor de sortida.

Per exemple, un detector de spam, analitza l'històric de missatges, veient que funció pot representar, segons els paràmetres d'entrada que es defineixin (el remitent, si el destinatari és individual o part d'una llista, si l'assumpte conté determinats termes, etc.), l'assignació de l'etiqueta spam o no és spam. Una vegada definida aquesta funció, en introduir un nou missatge no etiquetat, l'algorisme és capaç d'assignar-li l'etiqueta correcta.

L'aprenentatge supervisat se sol usar en problemes de classificació, com a identificació de dígit, diagnòstics, o detecció de frau d'identitat. També s'usa en problemes de regressió, com a prediccions meteorològiques, d'expectativa de vida, de creixement, etc. Aquests dos tipus principals d'aprenentatge supervisat, classificació i regressió, es distingeixen pel tipus de variable objectiu. En els casos de classificació, és de tipus categòric, mentre que, en els casos de regressió, la variable objectiu és de tipus numèric.

Alguns dels més freqüents en aprenentatge supervisat:

1. Arbres de decisió
2. Classificació de Naïve Bayes
3. Regressió per mínims quadrats

4. Regressió Logística
5. Support Vector Machines (SVM)

### 3.2.2 Aprenentatge no supervisat

L'aprenentatge no supervisat té lloc quan no es disposa de dades etiquetades per a l'entrenament. Només coneixem les dades d'entrada, però no existeixen dades de sortida que corresponguin a un determinat input. Per tant, només podem descriure l'estructura de les dades, per intentar trobar algun tipus d'organització que simplifiqui l'anàlisi. Per això, tenen un caràcter exploratori.

Per exemple, les tasques de clustering, busquen agrupaments basats en similituds, però gens garanteix que aquestes tinguin algun significat o utilitat. A vegades, en explorar les dades sense un objectiu definit, es poden trobar correlacions esporàdiques curioses, però poc pràctiques.

L'aprenentatge no supervisat se sol usar en problemes de clustering, agrupaments de co-ocurrència i profiling. Tanmateix, els problemes que impliquen tasques de trobar similitud, predicció d'enllaços o reducció de dades, poden ser supervisats o no.

Els algorismes no supervisats es poden usar per diferents tasques. Una tasca força clàssica és la de trobar grups dins a les dades. Podem usar algorismes de clustering que ens permeten associar a dades uns representants (centroïdes dels clústers). Això pot ser útil, per exemple, per a trobar grups de usuaris amb característiques comuns. Una altra tasca diferent és la de generar dades. En aquest treball ens centrem en usar models que permeten generar dades a partir de unes dades de entrenament, no en trobar grups a les dades.

## 3.3 Plantejament de l'algoritme

Un cop ja tenim els coneixements bàsics del machine learning necessitem pensar en quin algoritme necessitem per tal de poder aplicar-lo al nostre problema i desenvolupar-lo correctament.

Pel nostre problema disposarem de dades d'entrada no etiquetades, que seran les graelles de cada estil. És a dir, les dades d'entrada seran vectors que definiran graelles així un estil pot tenir diferents graelles.

D'aquesta forma el nostre algoritme ha de ser un algoritme de classificació no supervisat, ja que no disposem de cap tipus de dada etiquetada, només disposem de varies dades que sempre corresponen a un estil. Aquestes dades serien les features del nostre sistema.

Les features són variables individuals independents que actuen com a entrada al

sistema. Els models de predicció utilitzen funcions per fer prediccions. També es poden obtenir funcions noves a partir de funcions antigues mitjançant un mètode conegut com a enginyeria de característiques. Més simplement, podem considerar que una columna del conjunt de dades sigui una feature. De vegades, aquests també s'anomenen atributs. I la quantitat de funcions s'anomenen dimensions.

El que es vol aconseguir és que donat un estil ens generi una graella nova que segueixi les normes d'aquell estil que ha après la màquina. Així doncs necessitem un algoritme que ens permeti generar valors, no necessitem cap classificació com els algorismes anteriorment explicats.

Dintre del machine learning hi ha uns algorismes que ens permeten generar valor a partir dels valors d'entrada, aquests són els models generatius i discriminatius.

Els models generatius són un model per a generar valors aleatoris d'una dada observable, típicament donats alguns paràmetres ocults. Els models generadors contrasten amb els models discriminadors; un algoritme generador és un gran model probabilista de totes les variables, mentre que un algoritme discriminador proporciona un model sol per les variables etiquetades com a condicionals sobre les variables observades.

Per això un model generador pot ser utilitzat, per exemple, per simular (i.e. generar) valors de qualsevol variable en el model, mentre que un algoritme discriminador permet el mostreig únic de les variables condicionals. A la pràctica les dues classes són vistes com a complementàries o com diferents observacions del mateix procediment.

Així doncs, el que necessitem per al nostre problema és un algoritme generatiu que donats una sèrie de vectors que formen un estil, la màquina sigui capaç de generar una nova graella. Tot i això, aquest model ens provoca una restricció que és que haurem de tenir una màquina entrenada per a cada estil, ja que així aconseguim que els resultats siguin més precisos i correctes, ja que només permetrem que es generin graelles d'un estil per a cada màquina.

Tenim diferents tipus de models generatius. Els més coneguts són els següents:

- Model de barreja gaussiana i altres tipus de model de mescla
- Model ocult de Markov
- Gramàtica lliure de context probabilística
- Classificador bayesià ingenu
- Averaged one-dependence estimators
- Latent Dirichlet Allocation
- Restricted Boltzman machine

Després d'investigar entre totes les diferents opcions, vam arribar a la conclusió que la millor opció era utilitzar una Restricted Boltzmann machine (RBM) que ens permetria generar graelles a partir de les graelles ja generades anteriorment. Aquest tipus de màquina té algunes restriccions a l'hora d'entrar les dades, ja que han de seguir un format específic, a continuació explicarem com funciona aquest tipus d'algoritme i les seves característiques.

## 4 Boltzmann machine(BM) i Restricted Boltzmann machine (RBM)

A continuació explicarem d'una forma més detallada les màquines de Boltzmann i la seva variant, que és la que utilitzarem, la Restricted Boltzmann machine. Bàsicament ens centrarem en com entrenar i avaluar les màquines utilitzant el gradient estocàstic i analitzar les seves dificultats. També analitzarem els seus avantatges i inconvenients.

### 4.1 Boltzmann machine

La màquina Boltzmann (BM) és una xarxa neuronal recurrent estocàstica que consisteix en neurones binàries. La xarxa està totalment connectada i cada connexió entre dues neurones és simètrica, de manera que l'efecte d'una neurona en l'estat de l'altra és simètrica per a cada parell.

La probabilitat d'un estat particular  $X = [x_1, x_2, \dots, x_d]^T$  de la xarxa es defineix per l'energia de BM que es postula com

$$E(x|\theta) = - \sum_i \sum_{j>i} w_{ij} x_i x_j - \sum_i b_i x_i$$

on  $\theta$  indica paràmetres de la xarxa que consisteixen en una matriu de pes  $W = [w_{ij}]$  i un vector de biaix  $b = [b_i]$ .  $w_{ij}$  és el pes de les connexions sinàptiques entre les neurones  $i$  i  $j$ . La probabilitat d'un estat  $x$  es defineix com:

$$P(x|\theta) = \frac{1}{Z} \exp[-E(x|\theta)] \quad (4.1)$$

on

$$Z(\theta) = \sum_x \exp[-E(x|\theta)]$$

és la constant de normalització. Es dedueix de (4.1) que la probabilitat condicional d'una sola neurona que sigui 0 o 1 donada als estats de les altres neurones es pot escriure de la següent manera:

$$P(x_i = 1 | (x_{\setminus i}, W)) = \frac{1}{1 + \exp(-\sum_{j \neq i} w_{ij} x_j - b_i)} \quad (4.2)$$

on  $x_{\setminus i}$  denota un vector  $[x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d]^T$ .

Les neurones de BM solen estar dividides en visibles i ocultes  $x = [v^T, h^T]^T$ , on els estats  $v$  de les neurones visibles es corresponen a les dades observades, i els estats  $h$  de les neurones ocultes poden canviar lliurement. En cas de tenir neurones visibles i ocultes, es pot calcular la probabilitat d'una configuració específica de les neurones visibles marginalitzant sobre les neurones ocultes (és a dir, sumant sobre totes les possibles configuracions de neurones ocultes).

### 4.1.1 Entrenament Boltzmann machine

Els paràmetres de BM es poden obtenir a partir de les dades utilitzant l'estimació estàndard de màxima versemblança. Donat un conjunt de dades, la funció de likelihood es defineix com

$$\mathcal{L}(\theta) = \sum_{t=1}^N \log P(v^{(t)}|\theta) = \sum_{t=1}^N \log \sum_h P(v^{(t)}, h|\theta) \quad (4.3)$$

on les mostres  $v^{(t)}$  s'assumeixen que són independents entre si, i els estats  $h$  de les neurones ocultes han de ser marginalitzades.

El gradient de la probabilitat de registre s'obté prenent derivada parcial de  $\mathcal{L}(\theta)$  respecte als paràmetres  $w_{ij}$

$$\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{N}{2} [\langle x_i x_j \rangle_d - \langle x_i x_j \rangle_m]$$

on una notació abreviada  $\langle \cdot \rangle_{P(\cdot)}$  que denota la esperança calculada sobre la distribució de probabilitat  $P(\cdot)$ . A més, es van utilitzar  $d$  i  $m$  per designar dues distribucions de probabilitat  $P(h|\{v^{(t)}\}, \theta)$  i  $P(x|\theta)$ , respectivament. Són la probabilitat de les neurones ocultes quan les neurones visibles es subjecten a les mostres, i la probabilitat de totes les neurones sense cap neurona fixa.

Segons el signe de cada terme, els dos termes es poden anomenar la fase positiva i la fase negativa, respectivament. La fórmula d'actualització general d'un paràmetre  $w_{ij}$  és

$$w_{ij} \leftarrow w_{ij} + \eta [\langle x_i x_j \rangle_d - \langle x_i x_j \rangle_m] \quad (4.4)$$

on  $\eta$  denota la taxa d'aprenentatge.

Així doncs, a partir d'aquí, deixem que  $b$  sigui només un vector de biaixos  $b_i$  de les neurones visibles, i  $c$  sigui un vector dels biaixos de les neurones ocultes. A continuació, per als biaixos separats de les neurones visibles i ocultes, les regles d'actualització són, en analogia amb la regla d'actualització dels pesos,

$$b_i \leftarrow b_i + \eta [\langle v_i \rangle_d - \langle v_i \rangle_m] \quad (4.5)$$

i

$$c_j \leftarrow b_j + \eta [\langle h_j \rangle_d - \langle h_j \rangle_m] \quad (4.6)$$

on  $v_i$ ,  $h_j$  denoten neurones les neurones visibles i ocultes respectivament i  $b_i$  i  $c_j$  els biaixos de les neurones visibles i ocultes respectivament.

Tot i que les regles d'activació i aprenentatge de la BM estan clarament formulades, hi ha limitacions pràctiques en l'ús de la BM. Especialment, les fórmules d'actualització basades en gradients (4.4) - (4.6) no són computacionalment factibles, ja que les distribucions requerides tant en les fases positives com negatives només es poden obtenir després de computar la constant de normalització  $Z(\theta)$ .

---

**Algorithm 1** Passos de mostreig de Gibbs per al BM general

---

Dibuixa  $x_0$  uniformement des de l'espai de l'estat.

**repeat**

**for**  $i = 1 \dots d$  **do**

    Mostreig  $x_i$  fent l'equació (4.2)

**end for**

**until** que es reculli el nombre suficient de mostres, o el mostreig de Gibbs hagi arribat a l'equilibri.

---

La computació  $Z(\theta)$ , tanmateix, requereix fer un sumatori sobre un nombre exponencial de termes (les possibles configuracions de BM), i és senzillament impossible per a les BM "grans".

Un mètode per evitar la computació de la constant de normalització és utilitzar mètodes de Monte-Carlo que ens permeten aproximar integrals (o sumes de molts termes) per sumes amb un nombre factible de termes. A causa de la simplicitat de la regla d'activació d'una única neurona donada als estats d'altres neurones, un simple mostreig de Gibbs és suficient per obtenir gradients estocàstics.

El mostreig de Gibbs es pot implementar fàcilment perquè la distribució condicional de l'estat d'una sola neurona a BM donada als estats de totes les altres neurones ve donada per (4.2). Una descripció senzilla sobre com realitzar el mostreig de Gibbs amb BM es descriu en Algorisme 1.

Aquest enfocament pot reduir considerablement la càrrega computacional de les regles d'actualització de gradients. Si s'assumeix que la quantitat de mostres necessàries per explicar la distribució de probabilitat de tot l'espai estatal és prou menor que la mida de l'espai estatal, és a dir, el nombre de totes les combinacions possibles dels estats de les neurones, l'aprenentatge de BM ja no és computacional inviable.

Tanmateix, també existeixen altres tipus de limitacions en l'ús del mostreig de Gibbs per a la formació de BM. El major problema es deu a la connectivitat completa de BM. Atès que cada neurona està connectada i influïda per totes les altres neurones, es necessiten tants passos com la quantitat de neurones per obtenir una mostra de l'estat de BM. Fins i tot quan les neurones visibles es mantenen subjectes a les dades de formació, el nombre de passos necessaris per a una sola mostra



fresca encara és almenys el nombre de neurones ocultes. Això fa que les mostres successives de la cadena estiguin altament correlacionades entre elles i aquesta pobresa barreja afecta el rendiment de l'aprenentatge.

Una altra limitació d'aquest enfocament és que les distribucions multi modals són problemàtiques per al mostreig de Gibbs [7]. A causa de la naturalesa del mostreig de components, les mostres poden perdre alguns modes de distribució.

## 4.2 Restricted Boltzmann machine

Per superar les limitacions pràctiques imposades a la màquina general Boltzmann com el problema del mostreig ineficient, Smolensky va proposar una versió estructuralment restringida de la màquina Boltzmann anomenada Restricted Boltzmann Machine (RBM). Una RBM es construeix eliminant les connexions laterals entre les neurones visibles i les neurones ocultes. Per tant, una neurona visible només tindria vores connectades a les neurones ocultes, i una neurona oculta només tindria vores connectades a les neurones visibles. Ara, l'estructura d'una RBM es pot dividir en dues capes amb vores interconnectats. La relació entre BM i RBM s'il·lustra a la figura 9.

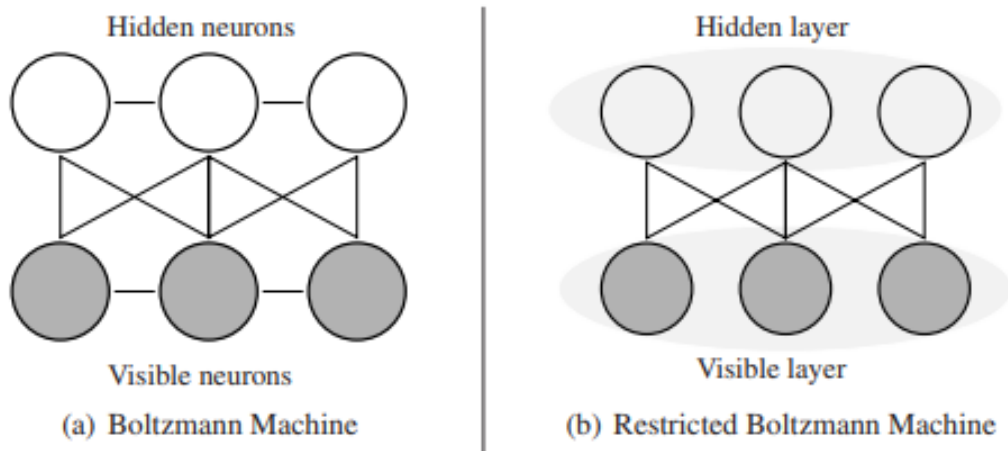


Figura 9: Il·lustració de la relació entre la màquina Boltzmann i la màquina restringida de Boltzmann.

Tot i que la restricció imposada podria suggerir que el poder de representació podria haver estat reduït, Le Roux and Bengio van demostrar que una RBM és una aproximació universal que pot modelar qualsevol distribució de probabilitat discreta [8].

A causa d'aquesta restricció, l'energia i la probabilitat de l'estat han de ser mo-

dificades en conseqüència a la següents fórmules:

$$E(v, h|\theta) = -v^T W h - b^T v - c^T h \quad (4.7)$$

$$P(v, h|\theta) = \frac{1}{Z(\theta)} \exp\{-E(v, h|\theta)\}$$

on ara els paràmetres  $\theta = (W, b, c)$  inclouen els biaixos  $b$  i  $c$ .

Atès que cada neurona oculta és independent una de l'altra donada totes les neurones visibles, és possible extreure explícitament les neurones ocultes i obtenir la probabilitat no normalitzada de les neurones visibles. La probabilitat d'un estat de neurones visibles  $v$  és, doncs,

$$P(v|\theta) = \frac{1}{Z(\theta)} \exp(b^T v) \prod_{j=1}^{n_h} (1 + \exp(c_j + \sum_{i=1}^{n_v} w_{ij} v_i)) \quad (4.8)$$

on  $n_v$  i  $n_h$  són el número de neurones visibles i invisibles respectivament.

#### 4.2.1 Entrenament Restricted Boltzmann machine

D'aquesta forma les regles d'aprenentatge d'una RBM passen a ser:

$$w_{ij} \leftarrow w_{ij} + \eta_w [\langle v_i h_j \rangle_d - \langle v_i h_j \rangle_m] \quad (4.9)$$

$$b_i \leftarrow b_i + \eta_b [\langle v_i \rangle_d - \langle v_i \rangle_m] \quad (4.10)$$

$$c_j \leftarrow c_j + \eta_c [\langle v_i \rangle_d - \langle v_i \rangle_m] \quad (4.11)$$

on s'utilitza la mateixa notació  $\langle \cdot \rangle_{P(\cdot)}$  que s'ha usat anteriorment.

Atès que una RBM és un cas especial de BM, és possible utilitzar el mateix mostreig de Gibbs durant l'aprenentatge. Gràcies a la seva estructura restringida, el mostreig de Gibbs es pot utilitzar de manera més eficient ja que les neurones de una capa sabent els valors de l'altra són independents (veure Figura 10).

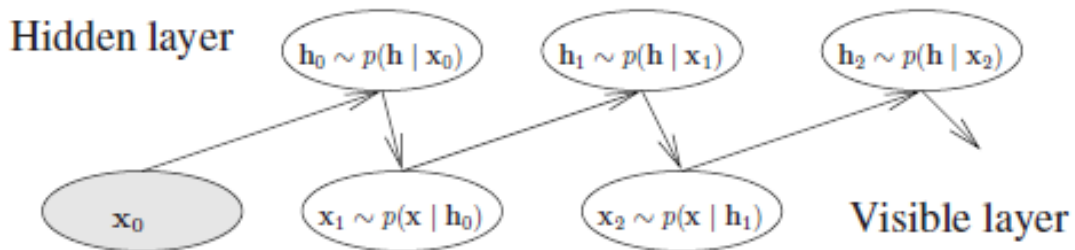


Figura 10: Visualització de la idea de com es realitza el mostreig de Gibbs en capes RBM.

Tanmateix, a mesura que augmenta el nombre de neurones en una RBM, el mostreig de Gibbs hauria de recollir un nombre més gran de mostres per explicar correctament la distribució de probabilitat que representa la RBM.

S'han proposat molts enfocaments per superar aquestes dificultats. Un enfocament popular és l'aprenentatge de Contrastive Divergence (CD) com un mètode aproximat per a la formació de la fase negativa de la RBM. L'aprenentatge de CD aproxima el calcul del gradient reemplaçant l'esperança sobre  $P(v, h|\theta)$  amb una mostra que s'obté executant  $n$  passos de Gibbs. La figura 11 il·lustra les distribucions  $P_0$  i  $P_n$ .

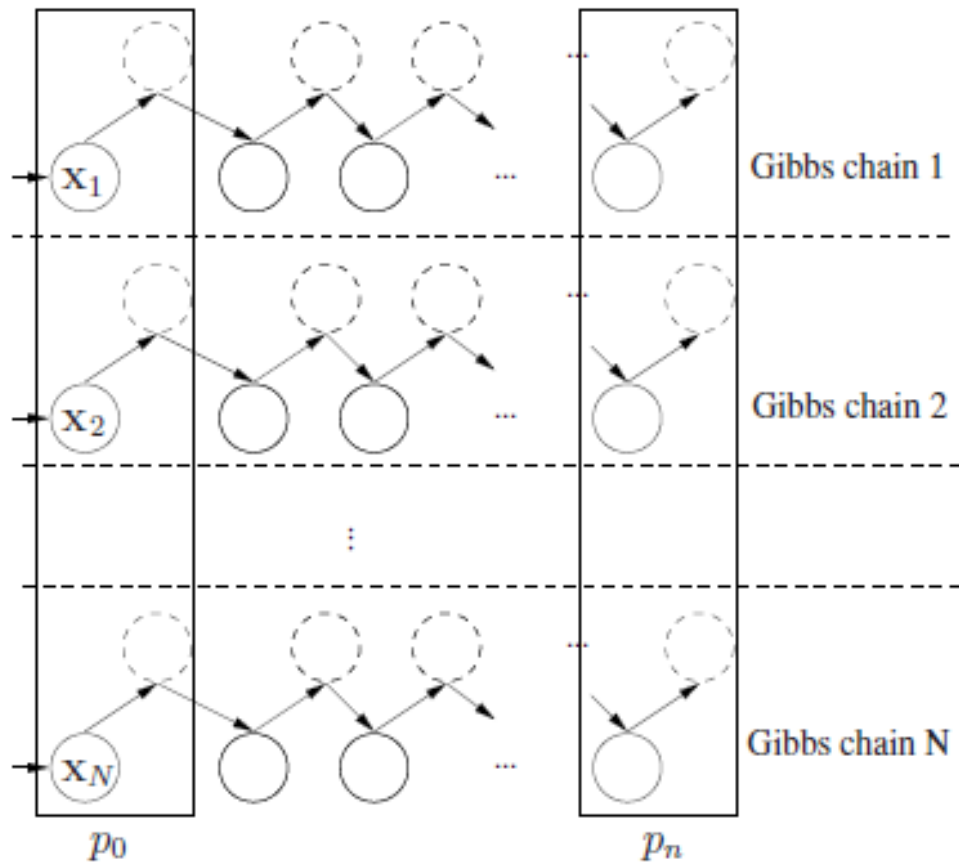


Figura 11: Visualització de com l'aprenentatge del CD obté la distribució empírica utilitzada en la fase positiva i la distribució del model aproximat utilitzada en la fase negativa. A la figura, cada fila representa la cadena de mostreig de Gibbs a partir de cada mostra de dades de formació, i  $P_0$  i  $P_n$  indiquen la distribució empírica i la distribució del model aproximat, respectivament.

Per als pesos, la fórmula d'aprenentatge dels CD, llavors, es defineix com

$$w_{ij} \leftarrow w_{ij} + \eta[\langle x_i h_j \rangle_{P_0} - \langle x_i h_j \rangle_{P_n}] \quad (4.12)$$

Cal assenyalar que el cas  $n = 0$  produeix la distribució empírica  $P(h|\{v^{(t)}\}, \theta)$  utilitzada en la fase positiva, mentre que el cas  $n = \infty$  produeix la veritable distribució de la fase  $P(x|\theta)$  [9].

Com que la direcció del gradient aproximat no és la del gradient exacte, se sap que l'aprenentatge de CD té alguns problemes [9]. No obstant això, s'ha demostrat que l'aprenentatge en CD funciona bé a la pràctica. EL problema de CD és que les mostres de  $P_n$  no necessàriament exploren tot l'espai d'estats. Per tant al no explorar algunes modes de la distribució el model pot acabar quedant esbiaixat. Aquest problema es mostra a la figura 12.

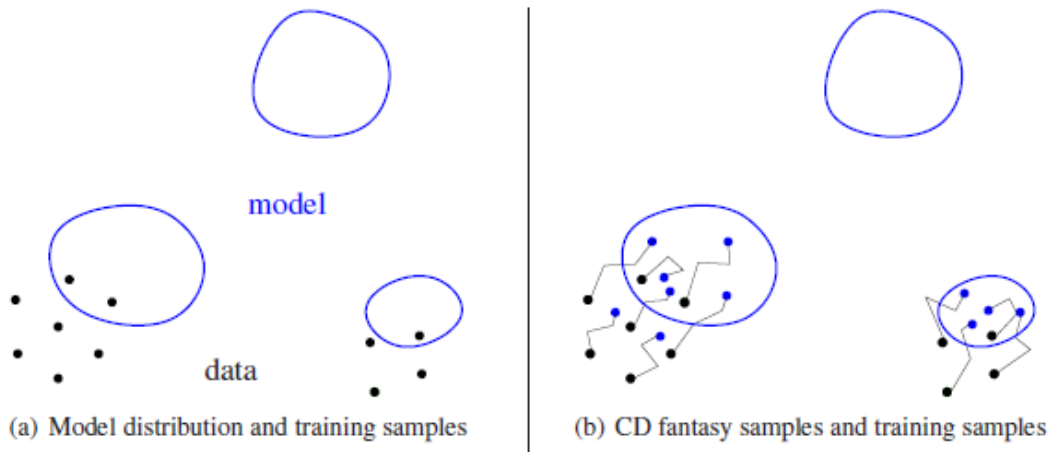


Figura 12: La figura esquerra mostra la distribució del model (blau) i les mostres d'entrenament (punts negres). Els punts blaus en la figura correcta indiquen les partícules de fantasia obtingudes mitjançant l'aprenentatge de CD. És evident que les partícules de fantasia no van explicar tot l'espai en perdre el mode a la part superior.

Per superar aquest problema, s'han proposat diferents enfocaments basats en l'aprenentatge de CD. Entre ells, Persistent Contrastive Divergence (PCD) que és l'extensió més senzilla de CD.

A cada pas d'actualització de gradients, l'aprenentatge de CD realitza el mostreig de Gibbs a partir de les mostres de dades de formació, mentre que l'aprenentatge de PCD comença el mostreig a partir de mostres model obtingudes en l'última actualització. D'aquesta manera, s'espera que les mostres del model exploren les modes de la distribució del model que no són pròxims a les mostres d'entrenament.

## 4.2.2 Avaluació d'una RBM

A continuació explicarem les formes d'avaluar una màquina de Boltzmann restrictiva. Tenim tres formes diferents per realitzar l'avaluació.

- Avaluació per probabilitat i mostreig rellevant d'importància
- Precisió de la classificació i altres mesures
- Visualització directe i inspecció de paràmetres

### 4.2.2.1 Avaluació per probabilitat i mostreig rellevant d'importància

Una forma natural d'avaluar el rendiment d'una RBM entrenada és calcular la probabilitat del model i les probabilitats de realitzar mostres de dades de prova sota la RBM formada. A més, la probabilitat de les mostres de dades aleatòries també es pot utilitzar com a mesura de la bondat de la RBM.

A causa de la restricció estructural, el resum explícit de les neurones ocultes (hidden) és força senzill (vegeu l'equació (4.8)), però, malauradament, la probabilitat d'una observació encara és intractable a causa de la constant de normalització. La constant de normalització només es pot computar exactament sumant exponencialment molts termes, i tret que la RBM sigui molt petita, és senzillament impossible. Per tant, en lloc de computar-ho exactament, cal utilitzar un mètode aproximat.

Per a l'estimació de la constant de normalització, aquesta tesi utilitza mostres d'importància recomanada (Annealed importance sampling, AIS) que s'ha utilitzat amb èxit per calcular la constant de normalització de la RBM [7].

L'AIS es basa en un mètode de mostreig d'importància simple (SIS) que podria estimar la proporció de dues constants de normalització. Per a dues densitats de probabilitat  $P_A(x) = \frac{P_A^*(x)}{Z_A}$  i  $P_B(x) = \frac{P_B^*(x)}{Z_B}$ , la relació de dues constants de normalització  $Z_A$  i  $Z_B$  es pot estimar mitjançant un mètode de mostreig de Monte Carlo sense cap biaix si és possible provar de  $P_A(\cdot)$ :

$$\frac{Z_B}{Z_A} = E_{P_A} \left[ \frac{P_B^*(x)}{P_A^*(x)} \right] \approx \frac{1}{M} \sum_{i=1}^M \frac{P_B^*(x_i)}{P_A^*(x_i)} \quad (4.13)$$

on  $X_i$  són mostres de  $P_A(X)$ . La qualitat de l'aproximació en termes de la variància depèn molt de la proximitat de  $P_A(\cdot)$  i  $P_B(\cdot)$ . Si  $P_A(\cdot)$  no és una aproximació gairebé perfecta a  $P_B$ , llavors la variància de l'estimació pot ser tan gran com l'infinit. Basat en SIS, AIS calcula la constant de normalització de la distribució del model calculant la ràtio de les constants de normalització de distribucions intermediàries consecutives que van des de l'anomenada distribució bàsica i la distribució de destinació.

La distribució base es tria de tal manera que la seva constant normativa  $Z_0$  es

pot calcular exactament i és possible recopilar mostres independents. Una selecció natural de la distribució base per a una RBM és una RBM amb zero pesos  $W$ . Això produeix la constant de normalització

$$Z_0 = \prod_i (1 + \exp\{b_i\}) \prod_j (1 + \exp\{c_j\}) \quad (4.14)$$

on els índexs  $i$  i  $j$  passen per totes les neurones visibles i ocultes, respectivament.

Mitjançant la computació del producte de les proporcions estimades de les constants de normalització intermèdia i  $Z_0$ , es pot estimar la constant de normalització de la RBM objectiu. L'algoritme que implementa AIS es descriu en Algorisme 2.

L'algoritme presentat descriu la construcció d'una RBM intermèdia seguint el que va proposar Salakhutdinov [7]. La distribució base està representada per una RBM amb pesos nuls, però els biaixos són idèntics als de la RBM objectiu. No obstant això, cal tenir en compte que hi ha altres possibilitats per construir distribucions intermèdies i triar una distribució bàsica. Per exemple, en els capítols següents, la distribució base és una RBM amb dos pesos zero i biaixos zero, de manera que no cal que cada RBM intermèdia mantingui el doble de neurones ocultes que la RBM objectiu.

---

**Algorithm 2** Estimació de la normalització constant per mostreig d'importància recomanada

---

Crear una seqüència de temperatures  $T_k$  tal que  $0 = T_0 < T_1 < \dots < T_K = 1$

Crear una RBM base,  $R_0$  amb paràmetres  $\theta_0 = (W_0, b, c)$ , on  $W_0 = 0$

Crear una seqüència de RBMs intermedis  $R_k$  tal que

- Té el doble de nodes ocults que la RBM objectiu té
- Els seus paràmetres són  $\theta_k = ((1 - T_k)W_0T_kW), [(1 - T_k)b_0T_kb], [(1 - T_k)c_0^T T_k c^T]^T$

**for**  $m = 1 \dots M$  **do**

Obté  $x_1$  de  $R_0$

**for**  $k = 1 \dots K - 1$  **do**

Mostreja  $x_{k+1}$  de  $R_k$  mitjançant un mostreig de Gibbs d'un pas a partir de  $x_k$ .

**end for**

Estableix  $u_m = \prod_{k=1}^K \frac{P_k^*(x_k)}{P_{k-1}^*(x_k)}$ , on  $P_k^*(\cdot)$  és una funció de distribució marginal no normalitzada de  $R_k$ .

**end for**

La estimació de  $\frac{Z_K}{Z_0}$  és  $\frac{1}{M} \sum_{m=1}^M u_m$ .

---

#### 4.2.2.2 Precisió de la classificació i altres mesures

És evident que a partir dels treballs de recerca esmentats anteriorment utilitzant xarxes neuronals profundes construïdes a partir de la pila de la RBM, les probabilitats d'activació oculta de la RBM entrenades en el conjunt de dades podrien millorar la precisió de classificació en comparació amb classificar el conjunt de dades en funció de les seves característiques (o features) originals.

Treballs recents suggereixen que les probabilitats d'activació oculta de RBM entrenada sense supervisió també ajuden a la tasca de classificació. Krizhevsky va utilitzar amb èxit un Gauss-Bernoulli RBM per extreure característiques d'imatges que ajuden a obtenir una alta precisió de classificació. A més, es va demostrar que les formes més sofisticades de RBM introduïdes recentment [10] van poder extreure característiques que són més útils per a la tasca de classificació.

A més, Coates et al. va mostrar que les característiques extretes pels models probabilístics apreses d'una manera no supervisada superen les contraparts supervisades, com ara xarxes neuronals convolucionals i la xarxa de conviccions profundes. Per tant, és convenient utilitzar la precisió de classificació de la RBM format com a mesura de rendiment.

Addicionalment, gràcies a la seva estructura bipartida i al mostreig de Gibbs capaç, l'error de reconstrucció també es pot utilitzar com a mesura per a l'avaluació del rendiment. Es defineix un error de reconstrucció

$$\mathcal{E}(X) = \|x - x_1\|_2$$

Tanmateix, aquestes mesures no reflecteixen directament la veritable qualitat de la RBM, ja que l'entrenament no maximitza ni minimitza cap d'aquestes mesures. Per tant, per a la resta d'aquesta tesi, els experiments valoren principalment la RBM format per la probabilitat del model i les probabilitats de les mostres de prova donades el model.

#### 4.2.2.3 Visualització directe i inspecció de paràmetres.

Una manera d'analitzar la qualitat d'un model entrenat és mirar les característiques (les ponderacions  $w_{ij}$ ) i els termes de polarització  $c_j$  corresponents a diferents neurones ocultes de la RBM modelat. Especialment ajuda quan es formen mostres de dades consisteixen en imatges que es poden visualitzar fàcilment.

Per exemple, es poden visualitzar funcions de la RBM formats en dígitos manuscrits. Cada característica o filtre s'assembla a una part de dígitos o una combinació de parts de dígitos. Quan l'aprenentatge falla, és fàcil observar característiques degenerades que són sorolloses funcions globals.

En cas de biaixos ocults, els valors mateixos suggereixen si cada neurona oculta

contribueix a la capacitat de modelització de la RBM. Les neurones que tenen un gran biaix  $c_j$  són la major part del temps actiu, i no són molt útils, ja que els pesos associats a ells es poden incorporar al biaix  $b$ . D'altra banda, les neurones ocultes que són majoritàriament inactives (p.Ex., Amb grans biaixos negatius  $c_j$ ) o les activacions dels quals són independents de les dades també són inútils, ja que la capacitat d'aprenentatge de la RBM no canvia encara que s'eliminin.

Igual que altres mesures indirectes presentades anteriorment, la visualització i la inspecció dels valors dels paràmetres s'haurien de realitzar acuradament. No hi ha cap mesura objectiva per la qualitat de les funcions visualitzades, i les característiques visualitzades i els valors dels biaixos poden evolucionar lentament a través de l'entrenament.

### 4.2.3 Dificultats i solucions convencionals

El fet que la funció de destinació no es pugui calcular exactament durant l'aprenentatge dificulta l'entrenament de les RBM. És computacional inviable saber quan l'aprenentatge ha convergit, o fins i tot no és fàcil dir si l'aprenentatge està passant realment. A més, en una RBM no és possible computar el gradient exacte de la funció que volem optimitzar. Per això recorrem a mètodes de Markov Chain Monte Carlo, per a generar una aproximació al gradient cada cop que volem canviar els pesos.

Un dels problemes de les RBMs és que no hi ha una única manera estàndard de entrenar-les. Els mètodes aproximats presentats en els apartats anteriors s'han demostrat que difereixen, especialment si els paràmetres d'aprenentatge no s'han triat de manera adequada. L'ús d'un millor mètode de mostreig MCMC, (com el "temperament paral·lel") s'ha demostrat que evita millor el comportament divergent.

Tot i els problemes explicats anteriorment, s'ha demostrat que una RBM és un aproximador universal de manera que amb suficients de neurones ocultes pot modelar qualsevol distribució de probabilitat discreta.

Tanmateix, a la pràctica, el nombre de neurones ocultes sempre és limitat i, dependent dels procediments d'aprenentatge, no totes les neurones ocultes contribueixen al poder de representació de la RBM.

Per exemple, aquestes neurones ocultes sempre actives no tenen sentit, ja que els pesos associats a ells es poden incorporar a biaixos. A més, qualsevol neurona oculta inactiva sempre no té sentit, des de llavors, l'eliminació de la neurona oculta no afecta la capacitat de modelatge de la RBM.

Idealment, cada neurona oculta hauria de representar una característica "significativa" diferent, per exemple, una part típica de la imatge. Tanmateix, ens hem adonat que sovint les neurones ocultes tendeixen a aprendre característiques que s'assemblen al biaix visible  $b$ . Aquest efecte és més destacat en la fase inicial d'a-



prementatge i en el conjunt de dades en què els bits visibles són majoritàriament actius, com ara 1-MNIST on es va desplegar cada bit de conjunt de dades manuscrits MNIST.

Es va obtenir una RBM amb 36 neurones ocultes en 1-MNIST, que és un conjunt de dades molt dens en comparació amb el MNIST original. 18 neurones ocultes no van poder aprendre cap característica útil, i són majoritàriament inactives.

Les altres 18 neurones són, en la seva majoria, actives i, tal com s'esperava, s'han après unes característiques globals que s'assemblen una mica al biaix visible.

Fins i tot quan les dades de formació no són denses, amb la petita quantitat de neurones ocultades, l'elecció inapropiada dels paràmetres d'aprenentatge i l'elecció inapropiada d'inicialització dels paràmetres, moltes neurones ocultes seran inútils. La visualització dels filtres apresos per una RBM amb 36 neurones ocultes entrenades a MNIST amb la taxa d'aprenentatge constant 0.1 i els pesos inicials mostrats de la distribució uniforme entre -1 i 1 són. No obstant això, encara existeixen aquestes neurones que són majoritàriament actives o majoritàriament inactives.

### 4.3 Algorisme seleccionat a implementar en codi

Un cop ja sabem quin algorisme hem d'utilitzar per tal de poder desenvolupar el nostre problema, ara hem de decidir com serà implementat, és a dir, si crearem nosaltres l'algorisme des de zero o utilitzarem alguna llibreria o algorisme ja conegut.

En aquest cas i després de fer molta investigació, en primer lloc es va decidir utilitzar el llenguatge de programació Python per tal de desenvolupar les parts essencials de l'algorisme. Python ens ofereix una gran quantitat de recursos per al machine learning i actualment juntament amb Scala són els principals llenguatges utilitzats per al desenvolupament de màquines.

Així doncs un cop escollit el llenguatge ara falta decidir si implementar una RBM des de 0 o utilitzar alguna llibreria que ens permeti usar una implementació ja desenvolupada.

Crear una màquina des de zero suposa un gran esforç i treball i pot suposar que la màquina no estigui ben dissenyada o que els algorismes aplicats i fórmules siguin incorrectes. Així doncs es va decidir utilitzar una llibreria la qual ens permetia tenir una RBM ben creada i amb les característiques que buscàvem. La màquina que es va decidir utilitzar seria de la llibreria *sklearn* la qual ens proporcionaria una màquina RBM anomenada *BernoulliRBM*[5].

Aquesta màquina es tracta d'una Restricted Boltzmann machine amb unitats visibles binàries i unitats ocultes binàries. Els paràmetres es calculen utilitzant la Probabilitat màxima estocàstica (SML), també coneguda com Persistent Contrastive Divergence (PCD). La complexitat del temps d'aquesta implementació és de

$O(d * 2)$  si assumim que el nombre de features és igual al nombre de components.

A més aquesta implementació ens permet utilitzar diferents mètodes i funcions per tal d'obtenir resultats i/o entrenar la màquina. Aquest són els mètodes i funcions principals d'aquesta implementació:

- *fit*( $X[, y]$ ): ens permet entrenar la màquina a partir de les dades  $X$ .
- *fit\_transform*( $X[, y]$ ): ens permet entrenar la màquina a partir de les dades  $X$  i després transformar-les.
- *get\_params*(*[deep]*): obtenim els paràmetres del model que hem generat.
- *gibbs*( $v$ ): s'executa un mostreig de Gibbs.
- *partial\_fit*( $X[, y]$ ): ens permet entrenar la màquina a partir de les dades  $X$  que han de contenir una part segmentada de les dades.
- *score\_samples*( $X$ ): computació de la probabilitat de  $X$ .
- *set\_params*(*\*\* params*): estableix els paràmetres del model.
- *transform*( $X$ ): computació de les probabilitats d'activació de la capa oculta.

Un exemple del ús d'aquesta màquina es troba en l'obtenció de features per la classificació de dígit. A continuació mostrarem un exemple de la creació d'una RBM utilitzant les funcions mencionades anteriorment.

```
X = np.array([[0, 0, 0], [0, 1, 1], [1, 0, 1], [1, 1, 1]])
model = BernoulliRBM(n_components=2)
model.fit(X)
#Sortida del resultat per la consola
BernoulliRBM(batch_size=10, learning_rate=0.1,
n_components=2, n_iter=10, random_state=None, verbose=0)
```

Com podem observar creem una llista de diferents elements els quals tots aquests són binaris, ja que la màquina només pot rebre entrades binàries. A continuació creem la màquina i li assignem el nombre de components i l'entrenem amb les dades d'entrada. A partir d'aquí podríem aplicar els altres mètodes per obtenir resultats o analitzar el seu funcionament.

Com podem observar al crear la màquina podem especificar-li alguns paràmetres d'entrada. Aquests, definiran el seu comportament. Aquests paràmetres són els següents:

- *n\_components*: Estableix el nombre de unitats binàries ocultes.
- *learning\_rate*: La velocitat d'aprenentatge de les actualitzacions de pes. Es recomana sintonitzar aquest hiperparametre. Els valors raonables es troben en el rang de  $10^{**} [0., -3.]$ .

- *batch\_size*: Nombre d'exemples per minibatch.
- *n\_iter*: Nombre d'iteracions sobre el conjunt de dades de formació que es realitzarà durant la formació.
- *verbose*: El nivell de verbositat. El valor predeterminat, zero, significa el mode silenciós.
- *random\_state*: Una instància de generador de números aleatoris per definir l'estat del generador de permutacions aleatòries. Si es dóna un enter, corregeix la llavor. Valors predeterminats del generador de nombres aleatoris global de *numpy*.

Aquests paràmetres no són obligatoris, es pot crear una màquina sense la necessitat de definir tots els paràmetres. En cas de no definir-ne cap, s'aplicaran els valors predeterminats.

Un cop tenim la implementació que usarem ara només ens queda preparar les dades per la entrada a la màquina. En els següents apartats anirem explicant el desenvolupament de la màquina creada sobre les dades que li proporcionem i el seu funcionament.

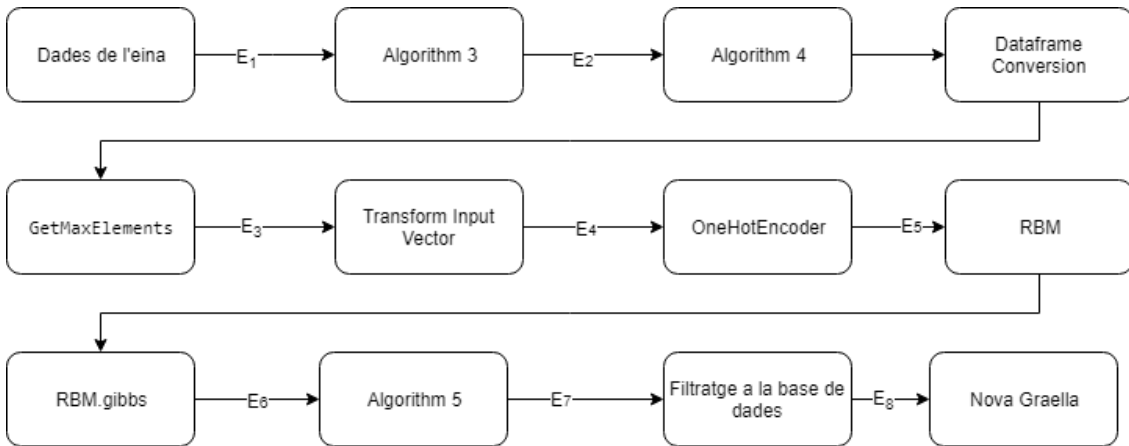
## 5 Desenvolupament de l'algoritme aplicat al problema

Un cop ja tenim tots els elements que componen el nostre problema només ens falta explicar el desenvolupament d'aquest utilitzant les dades que ens proporcionen i com utilitzem aquestes per a crear una màquina i generar resultats.

A continuació analitzarem les dades que rebem de l'aplicació, com es transformen per tal de ser aptes per la màquina, com funciona i com definim la màquina i com són les dades de sortida i com les transformem per tal de poder-les tornar a llegir amb l'aplicació.

### 5.1 Esquema del model

Tot seguit descriurem a alt nivell tots els passos pels quals passa una dada de entrada abans de arribar al model.



- $E_1 =$  Imatges col·locades a la graella.
- $E_2 = [id_1, id_2, \dots, id_{48}]$
- $E_3 = [[id_1^1, id_{2,i}^1, d_3^1, \dots, id_{255}^1], [id_1^2, id_{2,i}^2, d_3^2, \dots, id_{255}^2], \dots, [id_1^{48}, id_{2,i}^{48}, d_3^{48}, \dots, id_{255}^{48}]]$
- $E_4 = [[id_1^1, id_1^2, \dots, id_1^{48}], [id_2^1, id_2^2, \dots, id_2^{48}], \dots, [0, 0, \dots, id_y^{48}]]$
- $E_5 = [0_1, 0_2, \dots, 1_{10}, 0_{11}, \dots, 0_{262}]$
- $E_6 = [0_0, 0_1, \dots, 1_{125}, \dots, 0_{12576}]$
- $E_7 = [[id_1^1, id_{2,i}^1, d_3^1, \dots, id_{255}^1], [id_1^2, id_{2,i}^2, d_3^2, \dots, id_{255}^2], \dots, [id_1^{48}, id_{2,i}^{48}, d_3^{48}, \dots, id_{255}^{48}]]$
- $E_8 = [id_1, id_2, \dots, id_{48}]$

## 5.2 Dades obtingudes de l'aplicació

El primer pas com hem mencionat anteriorment és la generació de moodboards a partir de l'eina creada. Les graelles resultants són imatges formades a partir d'altres imatges, són representacions visuals que en aquest cas no podem entrar directament a la màquina ja que les hem de descompondre d'alguna manera per tal que les podem utilitzar per entrenar la nostra màquina.

Així doncs explicarem el format d'aquestes dades i quines dades hem obtingut en el nostre projecte.

### 5.2.1 Format de les dades

Com hem dit, cada imatge que forma un moodboard representa un objecte, en aquest cas un moble. Així doncs, cada moble ha de ser identificable d'alguna manera. La forma que tenim per identificar els mobles és un ID que ens proporciona la base de dades per a cada un dels mobles disponibles que tenim, d'aquesta forma podem identificar els diferents mobles que componen una graella.

Aixo ens permet obtenir un diccionari o una llista dels diferents ID que formen un moodboard el qual podré ser exportar d'una forma senzilla en un fitxer. Per tal d'obtenir aquesta llista només necessitem recórrer el moodboard, començant des de la primera cel·la, i obtenir els ID, que es troben internament en l'aplicació, d'aquells objectes col·locats. En cas que no hi hagi cap objecte a la casella, posarem un 0 per tal d'indicar que allà no hi ha cap objecte posat. L'algorisme 3 ens mostra el procediment a seguir.

---

**Algorithm 3** Generació de una llista de IDs a partir dels elements col·locats al moodboard.

---

Crear una llista  $L$  tal que  $L = [e_1, e_2, \dots, e_{48}]$  on  $e_x, x \leq 48$  correspon als ID dels objectes.

Moodboard  $M$  format per una llista de cel·les tal que  $M = [C_1, C_2, \dots, C_{48}]$ .

**for**  $i = 1 \dots 48$  **do**

    Comprovar si la cel·la  $C_i$  conté una imatge

**if**  $C_i$  té una imatge **then**

        Obtenir  $C_i(ID)$  i afegir  $C_i(ID)$  a la llista  $L$

**else**

        Afegir 0 a la llista  $L$

**end if**

**end for**

$L$  conté tots els ID del moodboard generat

---

Un cop apliquem l'algorisme descrit anteriorment, obtindrem una llista de ID, que abans eren simples imatges i no ens proporcionaven res d'informació a part de la visual. Ara tenim una llista amb els identificadors els quals cada un d'ells és un objecte que s'ha col·locat al moodboard. Així doncs de moment hem passat d'una



---

**Algorithm 4** Generació d'un diccionari amb les característiques dels objectes a partir de la llista de IDs.

---

Crear un Diccionari  $D$  tal que  $D = \{ "style" : [e_1, e_2, \dots, e_{48}] \}$  on  $e_x, x \leq 48$  correspon a les cel·les del moodboard i on  $e_x = [idc_1, idc_2, \dots]$ .

Llista  $L$  conté tots els ID obtinguts prèviament del moodboard generat.

**for**  $i = 1 \dots 48$  **do**

Obtenir llista  $C$  de les característiques de l'element  $L_i$  mitjançant crides a la base de dades

Llista temporal  $LL$  contindrà els ID de les característiques per afegir-les al diccionari

**for**  $j = 1 \dots M = length(C)$  **do**

Afegir  $C_j$  a  $LL$

**end for**

Afegir  $LL$  a  $D_i$

**end for**

$D$  conté una llista de totes els cel·les les qual cadascuna conté una llista de ID de les característiques de cada objecte que es troba al cel·la.

---

En aquest punt disposem d'un diccionari el qual ens permet saber a cada cel·la quines són les característiques del objecte que estava col·locat. Però aquest no es el format definitiu, ja que tal i com es va explicar, la màquina necessita una matriu com a dada d'entrada i en el nostre cas tenim un diccionari format per llistes que cada una conté una llista com a element i que cada una té una longitud diferent i la màquina necessita que cada element de la matriu, encara que siguin llistes, tinguin la mateixa longitud.

El format ideal per a nosaltres seria el format que obtenim directament de l'eina però no ens serveix perquè no ens proporciona suficient informació per tant l'hem hagut de transformar per tal d'obtenir més informació però ara ja no tenim el format desitjat ja que tenim un diccionari en el qual tenim una llista i cada element una llista de les característiques dels objectes.

Així doncs necessitem trobar una forma d'obtenir el format que desitgem, una llista on cada element contingui un identificador que correspondrà a un identificador d'una característica.

Una forma de resoldre aquest problema seria descomponent cada graella del diccionari en múltiples graelles que continguin un identificador per element. L'únic problema a resoldre aquí seria el cas del que les longituds són diferents ja que no tots els objectes tenen el mateix nombre de característiques. Així doncs hem de trobar una forma de solucionar aquest problema.

La solució més senzilla i més ràpida d'aplicar és buscar l'objecte amb més nombre de característiques. Un cop tenim aquest valor, aquest ens definirà el nombre de graelles a descompondre el diccionari, és a dir, si suposem que l'objecte amb més característiques en té 10, això vol dir que el màxim nombre de graelles noves que

hem de obtenir és 10 per cada moodboard generat, ja que si en fem menys no hi hauran totes els característiques d'alguns objectes i perdrem informació.

Per tal d'obtenir aquest valor aplicarem un algorisme de cerca en tot el diccionari. Fins ara només hem suposat que el diccionari només conté un moodboard però en podem tenir varis, per tant, haurem de fer una recerca entre tots els elements del diccionari i anar guardant la longitud de cada cel·la fins a trobar la longitud màxima. Per dur a terme aquest algorisme, primer transformarem el nostre diccionari en una llista ben construïda.

Per tal de realitzar la transformació utilitzarem un *Dataframe* a partir del qual ens organitzarà els elements en forma de taula. D'aquesta taula agafarem les *values*, que correspondran a cada cel·la i les afegirem a una llista així obtindrem una llista on cada element serà una llista i aquesta tindrà com a elements les diferents característiques de cada objecte.

$$input = [[id_1^1, id_{2,i}^1, d_3^1, \dots, id_{255}^1], [id_1^2, id_{2,i}^2, d_3^2, \dots, id_{255}^2], \dots, [id_1^{48}, id_{2,i}^{48}, d_3^{48}, \dots, id_{255}^{48}]]$$

Un cop ja tenim la llista apunt només necessitem aplicar la següent funció:

```
def getMaxLengthForElements(vector):
    maxLength = 0
    for i in range(len(vector)):
        tmp = vector[i]
        for j in range(len(tmp)):
            if (len(tmp[j]) > maxLength):
                maxLength = len(tmp[j])
    return maxLength
```

Aquesta funció ens retornarà el numero màxim de graelles que hem de generar de nou per tal d'obtenir a cada posició de cada element un identificador diferent.

Per a realitzar aquesta transformació hem de tenir en compte varis factors. El primer de tots hem de saber el nombre de moodboards que hem generat ja que haurem d'iterar sobre cada moodboard generat per transform-lo correctament.

I el segon punt i el més important hem de tenir en compte com tractarem aquelles cel·les les quals, la seva longitud sigui menor que la longitud màxima. En aquests casos la solució plantejada és afegir un valor nul, és a dir, un valor que no porti informació, en aquest cas el valor 0.

Això provocara que el resultat contingui llistes en les quals molts elements són 0 però no ens importa ja que només ens centrarem en els identificadors que no són 0. La funció per obtenir aquesta transformació és la següent:





### 5.2.2 Dades obtingudes del projecte

Ara que ja sabem com és el format de les nostres dades explicarem quines són les dades que hem obtingut al utilitzar l'eina durant el procés de desenvolupament del projecte.

En primer lloc es va demanar als experts la creació d'alguns moodboards d'un estil concret. En aquest cas l'estil s'anomenava Popular Modern. Es van crear 8 moodboards totalment diferents.

Utilitzant l'eina es va obtenir un fitxer on hi havia els 8 moodboards transformats a llistes de identificadors. Seguidament com hem explicat anteriorment i aplicant els algorismes descrits es van obtenir els identificadors de les característiques dels objectes i es va crear un diccionari el qual cada element era un moodboard format de llistes de identificadors de les característiques de cada element. En aquest punt es va buscar la llargada màxima de característiques que va resultar ser 15. Així doncs sortirien 120 graelles noves ja que per cada moodboard, aquest es descompondria en 15 graelles on cada element, contindria un identificador.

Finalment, les dades resultants obtingudes després d'aplicar la primera transformació, estan formades per una matriu de mida 120 per 48 on cada element és una llista, una graella, amb un identificador d'una característica a cada posició. En aquest cas, les dades proporcionades eren poques ja que amb 120 elements no és possible entrenar una màquina.

Per solucionar aquest problema vam decidir duplicar algunes dades ja que en realitat era com generar un moodboard igual a un altre. Aplicar això, podria afectar als resultats ja que les dades duplicades poden afectar al comportament de la màquina però en aquest cas com que les dades eren insuficients es va d'haver aplicar aquesta solució. L'altre solució era demanar als experts que generessin 500 moodboards però per falta de temps i coordinació no va ser possible. Així doncs necessitàvem obtenir una mida de dades més gran que 5000, per tant, vam duplicar les dades 51 vegades i vam obtenir una matriu de mida 6120 per 48. Podíem haver escollit un altre valor, però es va creure correcte tenir aquest volum de dades. Ara ja teníem una mida raonable per entrenar una màquina i podíem procedir a realitzar els següents passos.

## 5.3 Dades d'entrament

Cada màquina necessita un format específic per tal de llegir correctament les dades i poder aprendre d'elles per tant necessitem transformar les dades actuals en dades llestes per la ingesta. A continuació explicarem com són aquestes dades i quines dades hem obtingut.

### 5.3.1 Format de les dades

Un cop ja tenim les dades d'entrada transformades necessitem comprovar que la màquina que crearem és capaç de llegir-les. En el nostre cas crearem una RBM i segons la descripció que hem explicat anteriorment, el format de les nostres dades és el correcte però com que utilitzarem una implementació d'una llibreria, aquesta màquina té una sèrie de restriccions sobre les dades d'entrenament per a la màquina. La més important és que les dades han de ser binàries, és a dir, cada element, cada ID ha de estar en forma de llista de binaris. Tot i que el format de les nostres dades podria funcionar en altres màquines que utilitzen el mateix algorisme, en el nostre cas necessitem transformar-les per tal que la nostra implementació funcioni correctament. En aquest punt, les dades de les quals disposem no es troben en aquest format per tant necessitem trobar alguna forma de convertir-les.

Les dades que tenim són dades categòriques, és a dir, són variables que contenen valors d'etiqueta, a cada característica li assignem un número. Així doncs, una solució al problema seria codificar cada un dels ID que tenim en binari però aquesta solució no seria òptima ja que hauríem de fer una màquina que pogués descodificar cada un d'ells per tal de poder entendre el valor que li estem entrant i inclús en alguns casos no podríem entrar certes dades si el número a codificar fos molt llarg.

La solució més òptima i més assequible és utilitzar els denominats *One Hot Vectors*. Aquests tipus de vectors són un grup de bits entre els quals les combinacions vàlides de valors són només aquelles amb un sol bit alt(1) i tots els altres valors amb bits baixos(0). És a dir, ens permet codificar qualsevol número en forma d'una llista que serà tant llarga com el número més alt i que contindrà tots els valors 0 excepte la posició en la qual representi el valor a codificar.

Un exemple fàcil per a veure el funcionament d'aquests vectors és aplicar al nostre cas. Disposem de 262 característiques diferents a la nostre base de dades per tant si volem codificar un identificador d'una característica concreta, per exemple l'identificador 10, el codificarem de la següent manera:

$$10 = [0_1, 0_2, \dots, 1_{10}, 0_{11}, \dots, 0_{262}]$$

D'aquesta forma podrem codificar tots els identificadors de forma binària. Per transformar cada identificador podem crear un algorisme que ens converteixi un número en un vector o també podem utilitzar eines que ens proporcionen algunes llibreries com és el cas de la llibreria *sklearn* que ens proporciona un codificador per codificar les nostres dades. Aquest codificador s'anomena *OneHotEncoder*[6].

Aquest codificador ens permet transformar les nostres dades d'entrada en una matriu, on cada columna correspon a un possible valor d'una característica. Al crear un codificador d'aquest tipus podem establir alguns paràmetres d'entrada per tal de definir com serà i com seran les dades que li entrem i com són les que volem obtenir.

De tots els paràmetres que podem definir els més importants són:

- *n\_values*: Nombre de valors per característica. En aquest cas el nombre de identificadors màxims que hi pot haver. Com hem dit anteriorment, aquest valor és 262.
- *sparse*: Boolea que ens indica si utilitzarem matrius disperses, *sparse matrix*, o matrius denses i el format del resultat.

A part d'aquests paràmetres també té mètodes que ens permetran realitzar certes accions. El mètodes més importants i que més endavant utilitzarem són els següents:

- *fit(X[, y])*: ens permet entrenar el codificador amb les dades X.
- *fit\_transform(X[, y])*: ens permet entrenar el codificador a partir de les dades X i després transformar-les.
- *transform(X)*: ens permet transformar unes dades X en un one hot vector.

Així doncs, un cop creat el codificador i entrenat amb les dades que hem obtingut de la primera transformació podem obtenir les nostres dades llestes per la ingesta de la màquina. La creació d'aquestes dades provocarà que la mida de les dades augmenti sobretot en les files de la matriu, en aquest cas, en la llista d'elements de cada cel·la ja que cada cel·la contindrà una successió de 0 i 1 i aquesta serà tant llarga com el nombre de valors per característica que li haguem definit.

### 5.3.2 Creació del codificador i dades obtingudes

Ara ja sabem què necessitem per transformar les nostres dades per realitzar la ingesta a la màquina i així entrenar-la. A continuació explicarem com hem creat el nostre codificador, quins paràmetres hem definit i quines són les dades obtingudes.

Recordem que de la primera transformació de les dades vam obtenir una matriu de mida 6120 per 48 que contenia les graelles amb els identificadors de les característiques dels objectes del moodboard. El primer que hem de definir és el nombre de valors per característica, en aquest cas disposem de 262 característiques diferents, el que és el mateix, que 262 identificadors diferents per tant, el paràmetre *n\_values* serà igual a 262.

La segona cosa i la que més ens afectarà alhora del rendiment del codificador és el booleà *sparse*. Com hem dit aquest valor ens indica quin tipus de matriu farem servir i obtindrem, si serà una *sparse matrix* o una matriu densa. Per entendre bé la decisió a prendre necessitem saber les diferències entre els dos tipus.

Les *sparse matrix* només emmagatzemen les entrades que no són zero, aquelles entrades de la matriu que tenen algun valor mentre que les denses emmagatzemen tots els valors. Així doncs la diferència es troba alhora de treballar amb elles, és a dir, a l'hora de recórrer les matrius o de comprimir-les. En aquests cas és molt millor treballar amb *sparse matrix* ja que el procés serà més ràpid i sobretot quan

es tracten de matrius molt grans. En cas contrari, si es tracten de matrius petites es poden utilitzar matriu denses.

Aplicant això al nostre cas podem fer servir les dos formes ja que la matriu és gran però no exageradament enorme. De tot formes a continuació mostrarem dos exemples, Figura 16, de la creació del nostre codificador i el resultat que dona, un utilitzant *sparse matrix* i l'altre utilitzant matrius denses.

```
%%time
oneHotEncoder = OneHotEncoder(262, sparse=False).fit(f_data)
oneHotData = oneHotEncoder.transform(f_data)
print(oneHotEncoder)
print(oneHotData.shape)

OneHotEncoder(categorical_features='all', dtype=<class 'numpy.float64'>,
              handle_unknown='error', n_values=262, sparse=False)
(6120, 12576)
Wall time: 341 ms
```

```
%%time
oneHotEncoder = OneHotEncoder(262, sparse=True).fit(f_data)
oneHotData = oneHotEncoder.transform(f_data)
print(oneHotEncoder)
print(oneHotData.shape)

OneHotEncoder(categorical_features='all', dtype=<class 'numpy.float64'>,
              handle_unknown='error', n_values=262, sparse=True)
(6120, 12576)
Wall time: 59.8 ms
```

Figura 16: Creació del codificador de les dos forma. A dalt amb matrius denses i a baix amb *sparse matrix*

Com podem observar, el resultat final és una matriu del mateix nombre de files que la matriu anterior però ara els elements són més grans ja que multipliquem cada element per 262. També podem observar que utilitzant *sparse matrix* el temps es redueix. En aquest cas el temps és molt petit però es pot observar una diferència. Així doncs el millor mètode és fer servir *sparse matrix*.

El contingut de la nostre matriu serà per a cada fila una successió de 0 i 1, al tractar-se d'una *sparse matrix* si la mostrem per pantalla només veurem aquelles posicions que contenen un valor diferent a 0. Ara ja tenim les dades llestes per entrar a la màquina i entrenar-la però primer de tot hem de crear-la d'acord a les dades que tenim.

## 5.4 Creació i entrenament de la RBM

Un cop ja tenim les dades transformades per l'entrenament de la màquina només ens queda crear la màquina i entrenar-la per tal d'obtenir resultats. A continuació explicarem la creació de la nostra màquina i analitzarem el seu comportament.

### 5.4.1 Creació de la RBM a partir de les dades

En apartats anterior hem vist com crear una RBM estàndard utilitzant la llibreria. Però ara ja tenim les dades i necessitem trobar els paràmetres adients per tal que la creació d'aquesta màquina sigui la correcta o la més adient per al nostre problema. Els principals camps que definirem seran el *n\_components*, *learning\_rate* i *n\_iter*.

Per començar hem de definir el nombre de unitats ocultes que tindrà la nostra màquina.

Hi ha varies formes de trobar el número perfecte de unitats ocultes aplicant fórmules però en el nostre cas hem decidit que aquest nombre serà el valor total de les característiques disponibles que en aquest cas es 262, així disposem d'una unitat oculta per a cada un dels valors de les dades.

Un cop ja tenim el nombre de components decidit hem de decidir la velocitat d'aprenentatge. Tal i com es va explicar anteriorment ha d'estar en el rang de  $10^{**}[0., -3.]$ , així doncs a partir de la documentació de la llibreria [5] i partir d'alguns exemples que també es trobaven a la documentació vam poder establir un valor relativament correcte per a la nostre màquina.

Es va establir un valor de 0.01. Aquest valor és una funció que el que defineix és com de ràpid canvien el pesos dels valors al llarg del aprenentatge. Si aquest valor és massa alt, és possible que es perdi el punt 0 del pendent i si es massa baix pot provocar que es tardi molt a arribar aquest punt.

Finalment queda definir el nombre d'iteracions. Ja es sap que en una màquina d'aquest tipus, cada iteració pot durar molt de temps per tant no volíem que tardes molt a realitzar totes les iteracions. En un principi vam establir unes 10 iteracions per tal de comprovar el temps i la seva eficiència. Finalment, a partir dels resultats es va establir un valor de 20 iteracions.

Així doncs, un cop ja tenim tots els valors definits només ens queda crear la màquina. A continuació, a la figura 17 es mostra com s'ha realitzat la creació de la màquina.

### 5.4.2 Entrenament de la RBM

Un cop ja creada la màquina necessitem entrenar-la per tal de, més endavant, poder generar noves dades i així poder analitzar el seu funcionament amb les dades que li

```
RBM_Machine = BernoulliRBM(n_components=262, learning_rate=0.01, n_iter=20, random_state=0, verbose=True)
print(RBM_Machine)

BernoulliRBM(batch_size=10, learning_rate=0.01, n_components=262, n_iter=20,
             random_state=0, verbose=True)
```

Figura 17: Creació de la RBM amb els paràmetres definits anteriorment

passem. Per tal d'entrenar-la només cal cridar el mètode *fit* de la màquina. Al cridar aquest mètode li hem de passar les dades les quals volem que siguin les utilitzades per l'entrenament. En aquest cas li hem de passar la *sparse matrix* que hem obtingut anteriorment.

A mesura que es vagin realitzant les iteracions es mostrarà per pantalla la iteració la qual s'acaba de realitzar, la *pseudo-likelihood* i el temps que ha tardat. La *pseudo-likelihood* és una estimació bastant decent sobre el rendiment del model actual encaixant les dades de formació. Com més baix sigui, millor. Així doncs al entrenar la nostra màquina creada obtenim un resultat com aquest:

```
RBM_Machine.fit(oneHotData)
[BernoulliRBM] Iteration 2, pseudo-likelihood = -78.52, time = 57.54s
[BernoulliRBM] Iteration 3, pseudo-likelihood = -80.80, time = 60.48s
[BernoulliRBM] Iteration 4, pseudo-likelihood = -74.14, time = 55.01s
[BernoulliRBM] Iteration 5, pseudo-likelihood = -77.79, time = 52.40s
[BernoulliRBM] Iteration 6, pseudo-likelihood = -75.31, time = 53.13s
[BernoulliRBM] Iteration 7, pseudo-likelihood = -74.91, time = 52.39s
[BernoulliRBM] Iteration 8, pseudo-likelihood = -77.24, time = 52.09s
[BernoulliRBM] Iteration 9, pseudo-likelihood = -75.42, time = 52.31s
[BernoulliRBM] Iteration 10, pseudo-likelihood = -75.18, time = 56.31s
[BernoulliRBM] Iteration 11, pseudo-likelihood = -81.44, time = 61.02s
[BernoulliRBM] Iteration 12, pseudo-likelihood = -73.93, time = 55.00s
[BernoulliRBM] Iteration 13, pseudo-likelihood = -71.03, time = 66.80s
[BernoulliRBM] Iteration 14, pseudo-likelihood = -79.37, time = 61.06s
[BernoulliRBM] Iteration 15, pseudo-likelihood = -76.83, time = 62.95s
[BernoulliRBM] Iteration 16, pseudo-likelihood = -73.65, time = 60.16s
[BernoulliRBM] Iteration 17, pseudo-likelihood = -76.19, time = 66.38s
[BernoulliRBM] Iteration 18, pseudo-likelihood = -76.36, time = 64.10s
[BernoulliRBM] Iteration 19, pseudo-likelihood = -73.28, time = 61.93s
[BernoulliRBM] Iteration 20, pseudo-likelihood = -75.75, time = 57.33s
Wall time: 19min 22s
```

Figura 18: Resultat del entrenament de la màquina

Com podem observar el temps de cada iteració es similar i la *pseudo-likelihood* va variant a mesura que passen les iteracions. Aquesta informació ens permet analitzar el comportament de la nostra màquina, és a dir, si volem canviar el valor la *pseudo-likelihood* per tal que sigui més baix i millor necessitem anar realitzant proves i anar canviar els diferents paràmetres de la màquina. En aquest cas, una

solució seria canviar el *batch\_size* però això comportaria més temps per iteració i a més tenint en compte que el nombre de dades actual és baix en comparació a altres màquina no es factible canviar aquest valor.

La qüestió final és fer proves fins a trobar una implementació que ens sigui assequible, que s'adapti a les nostres necessitats i a les del problema proposat. Per tal de trobar aquesta implementació es van realitzar diferents proves amb diferents valors però els valors de temps eren molt elevats i la *pseudo-likelihood* era molt alta en comparació a les altres implementacions.

Finalment un cop acaben totes les iteracions ja tenim la màquina creada per tal d'obtenir valors nous. Per tal d'evitar entrenar la màquina cada cop que volem generar valors nous es va decidir guardar tant les dades d'entrenament com la màquina en un format de fitxer *pickle*. Aquest format ens permet guardar la nostra màquina i les nostres dades sense la necessitat d'ocupar molt d'espai i també ens permet més endavant carregar de nou la màquina, ja entrenada, les nostres dades i els paràmetres tant de la màquina com del codificador.

## 5.5 Dades de sortida

Després d'haver entrenat la màquina només ens queda una pas i és generar moodboards. Però per tal de generar moodboards hem de saber com obtenir valors nous a partir de la màquina i en quin format es troben i quin format necessitem per visualitzar-los a la eina.

### 5.5.1 Dades obtingudes de la RBM

Un cop la màquina està entrenada necessitem generar dades noves. Per tal de generar dades noves, la llibreria amb la qual hem implementat la màquina ens proporciona un mètode per tal de generar aquestes dades. Aquest mètode es tracte del mostreig de Gibbs, *gibbs(v)*. El mostreig de Gibbs és un algoritme per generar una mostra aleatòria a partir de la distribució de probabilitat conjunta de dues o més variables aleatòries. A partir d'un vector de unitats visibles, de valors visibles i vàlids genera un altre vector d'unitats visibles mitjançant el mostreig de Gibbs.

Així doncs, aquesta funció ha de rebre com a paràmetre un *one hot vector*. En aquest cas, aquest vector ha de correspondre a una dada de les d'entrenament ja que són les aquestes són les úniques dades vàlides. En molts casos, es té unes dades d'entrenament i unes dades de test, però en aquest cas al tenir poques dades no s'ha dividit en dos tipus de dades.

D'aquesta forma, al passar un vector format de 0 i 1 a la funció *gibbs*, aquesta ens retornarà un vector de unitats visibles de la màquina.

Aquest vector que obtindrem no estarà format per 0 i 1 sinó que estarà format per booleans a cada posició. A més la seva llargada no serà de 48 elements com la



mida d'un moodboard sinó que la mida serà de 12576 com la mida de cada vector de la matriu d'entrada a la màquina. Així doncs el format d'aquest vector resultant serà:

$$[False_0, False_1, \dots, True_{125}, \dots, False_{12576}]$$

Això ens suposa alguns problemes ja que nosaltres volem un vector que tingui una mida de 48 i on a cada posició tingui un o més d'un identificador de característiques. El primer problema a observar és que haurem de transformar el vector obtingut per obtenir identificador i el segon problema és que a si només obtenim un vector només tindrem un identificador per cel·la o poder cap. Per tant haurem de realitzar el mostreig de Gibbs varis cops fins que tinguem un vector correcte.

Primer de tot hem de transformar el vector per obtenir un vector de 0 i 1. Per fer això només hem de recórrer el vector i substituir tots els valors *False* per 0 i els valors *True* per 1.

Un cop realitzat aquest procés ja tenim un vector amb el següent format:

$$[0_0, 0_1, \dots, 1_{125}, \dots, 0_{12576}]$$

El següent pas és convertir aquest vector en un de mida 48 i amb identificadors a cada posició on cada identificador pot tenir un valor entre 1 i 262 que són els identificadors de les característiques que tenim. També poden tenir un valor 0 el qual ens indica que en aquella posició no hi ha cap objecte. Per realitzar aquesta transformació aplicarem el següent algorisme:

---

**Algorithm 5** Generació d'un vector amb identificadors a cada posició a partir d'un vector de 0 i 1

---

Vector  $V$  tal que  $V = [i_0, i_1, \dots, i_{48}]$  on  $0 \leq i \leq 262$ .

Vector  $D$  tal que  $D = [j_0, j_1, \dots, j_{12576}]$  on  $j = 0$  o  $j = 1$ .

**for**  $i = 1 \dots 48$  **do**

Obtenir el vector  $K$  tal que  $K = D[(262 * i) : ((i + 1) * 262)]$

**if**  $K$  no conté 1 **then**

Afegir 0 a  $V$

**else**

Afegir  $K.index(1)$  a  $V$

**end if**

**end for**

$V$  conté a cada posició un *int* que és l'identificador d'una característica d'un objecte.

---

Així doncs, aplicant l'algorisme anterior obtindríem un vector amb els identificadors corresponents a les característiques. Però, tot i això encara ens falta solucionar el problema d'obtenir més d'un valor per cada element. Per tal d'aconseguir el format final correcte necessitem tenir més d'un identificador a cada element, és a dir, una llista de identificadors a cada posició tal i com obteníem les dades després de transformar-les al sortir de l'eina.





## 6 Resultats i conclusions i línies de continuació

Ara ja tenim totes les parts enllaçades i preparades per generar moodboards, l'últim pas que ens queda és obtenir resultats, avaluar-los, treure'n conclusions i explicar les possibles línies de continuació que existeixen per aquest projecte que hem desenvolupat.

### 6.1 Qualitat de les graelles generades

Avaluar models generatius com la màquina que hem creat és molt complicat ja que no tenim un resultat estàndard per avaluar el resultat obtingut i no qualsevol persona pot detectar si aquell resultat és bo o dolent. Així doncs la nostra forma d'avaluar els nostres resultats serà utilitzant l'ajuda d'un expert per a valorar els moodboards obtinguts.

L'avaluació que definirem per avaluar els moodboards generats serà una avaluació qualitativa. Es generaran 10 moodboards utilitzant la màquina i al eina respectivament i se'ls ensenyaran al expert i haurà de valorar per cada un d'ells quins són els objectes que fallen i perquè. És important emfatitzar que degut a dificultats tècniques sols disposàvem de 8 dades d'entrenament.

D'aquesta forma obtindrem els moodboards més ben generats i els que més objectes incorrectes tenen. Aquesta avaluació ens permetrà tenir una primera idea si la màquina esta generant els moodboards correctament, és a dir, si conté objectes correctes i si és el cas, dels que s'han generat quina és la mitja de elements correctes.

També ens permetrà saber les causes dels objectes que han fallat i així intentar millorar la màquina per tal que tals objectes no siguin generats o que valori uns més que els altres. A continuació es mostren els moodboard generats.

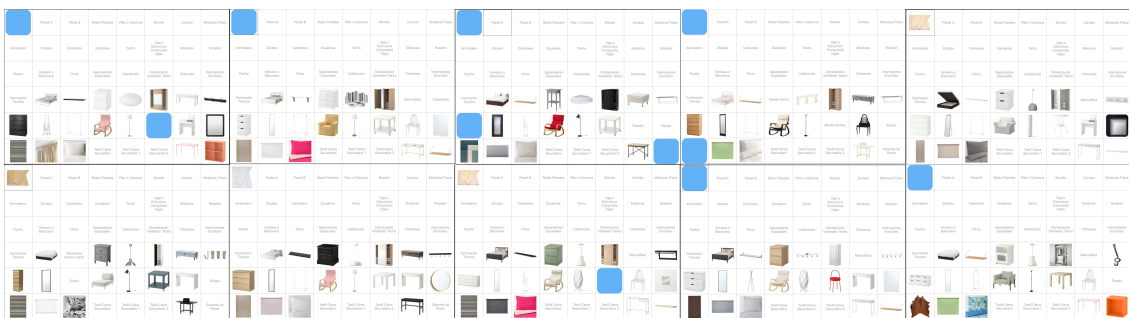


Figura 20: Resultat de la generació de 10 moodboards.

Com podem observar tots els moodboards segueixen un esquema similar, les caselles on no hi ha objectes sempre solen ser les mateixes. També cal observar que hi ha objectes que no tenen imatges, això pot complicar l'avaluació ja que no sabem



Com podem observar de les 10 imatges que s'han generat la majoria tenen algun element incorrecte o que no s'adequa al estil. L'expert a part de dir quins són els que estan malament també ens indica algunes de les raons per les quals aquests objectes són incorrectes o no s'adeqüen a l'estil.

En aquest cas les imatges corresponen a l'estil de *Popular Modern*. Així doncs, molts dels objectes que s'han col·locat no segueix aquest estil, aquesta línia de popular modern. La majoria són popular tradicional o tradicionals purament per tant, la col·locació d'aquests no es correcte a dins la imatge.

Si analitzem la quantitat d'elements de tots els moodboards, tenim que en els deu moodboards tenim un total de de 191 objectes col·locats, que serien una mitjana de 19 objectes per moodboard així doncs podem saber el tant per cent d'objectes correctes i incorrectes que hi ha. Així doncs tenim que el 35% dels objectes col·locats són incorrectes. A partir d'això podem dir que encara que tots els moodboards hagin marcats com a incorrectes, més de la meitat dels objectes col·locats són correctes amb l'estil seleccionat. Això ens indica que el funcionament de la màquina és correcte tot i que no del tot, no és perfecte, perquè si funcionés perfectament hauríem de tenir un percentatge casi nul d'elements incorrectes. Però podem afirmar que dintre del marge que tenim l'algorisme ha funcionat bastant bé ja que no hem passat del 50% d'objectes incorrectes.

### 6.1.1 Possibles problemes

- Les dades d'entrada són escasses. Només teníem 8 graelles generades correctament i tot i que amb les transformacions hem aconseguit augmentar el nombre de graelles i duplicant dades no són suficients ja que per a qualsevol problema d'aquest estil necessitem una mida de dades molt gran i amb dades no duplicades. Les dades que teníem per aconseguir els resultats ens han frenat, ens han provocat, que els resultats no fossin del tot correctes ni que fossin com esperàvem. L'obtenció de poques dades també es causada a conseqüència alienes al projecte, relacionades amb l'empresa i la gestió del projecte.
- El filtratge de les dades de sortida és incorrecte. Quan obtenim les dades de sortida de la màquina i les transformem per obtenir els identificadors de les característiques, realitzem crides a la base de dades per obtenir els identificadors dels objectes. Aquestes crides es fan unint els diferents identificadors utilitzant un operador *AND* però la base de dades ho interpreta com una suma, *OR*, així doncs ens retorna una llista d'objectes bastant més extensa del que ens esperàvem i per tant hem d'agafar un objecte aleatori. Si el funcionament fos el correcte ens retornaria una llista d'objectes però en aquest cas seria molt petita, de dos o tres objectes que compleixen totes les condicions.
- El format de les dades. Hem hagut de transformar les dades varis cops per tal que la màquina les pogués llegir. Aquestes transformacions, la majoria consistien en separar les graelles en múltiples elements. Això pot haver afectat

una mica al aprenentatge de la màquina. També afecte el fet anterior de la quantitat de dades, ja que també ens ha obligat a dur tantes transformacions al llarg del projecte.

- Els colors no concorden. La màquina no té en compte la relació de colors entre objectes, no estableix cap relació entre els objectes que es col·loquen ja que només s'analitzen les característiques d'aquests. Així doncs és normal que els colors dels objectes no tinguin relació quan en la generació prèvia de moodboards l'expert col·locava els objectes per tal de tenir coherència entre ells però alhora d'entrenar la màquina aquesta relació no s'analitza profundament.

## 6.2 Conclusions

Amb el desenvolupament d'aquest projecte hem après que la realització d'una tasca manual, repetitiva i costosa de temps, com és la generació de moodboards, es pot millorar aplicant tècniques de Intel·ligència Artificial.

També hem après com dissenyar una aplicació en unity que dona una interfície gràfica per tal de poder fer més fàcil la tasca de la generació de moodboards als dissenyadors.

La part més important ha sigut crear la màquina amb capacitat d'aprenentatge ja que hem hagut de aprendre el funcionament bàsic del machine learning per tal de decidir quin tipus de màquina era millor pel problema descrit. La creació de la màquina ens ha permès establir un sistema d'aprenentatge per tal de generar moodboards a partir dels creats anteriorment i ens ha permès automatitzar tota la tasca per tal d'evitar que els dissenyadors comencin a treballar des d'un moodboard buida.

Els resultats que hem obtingut després de realitzar totes les proves que hem cregut necessàries no han tingut la qualitat esperada. Tot i això, les resultats ens han permès donar una eina que pot servir per a donar idees de moodboards inicials als dissenyadors.

És important remarcar que la qualitat dels resultats pot ser deguda a la falta de entrenament que disposàvem. Per aquest projecte no hem pogut disposar de més dades per culpa de la falta de temps dels dissenyadors de la empresa que no ens van proporcionar un nombre de moodboards gaire elevat. En conseqüència l'entrenament de la màquina s'ha fet amb tant sols 8 dades (graelles) que són moltes menys de les que es fan servir per a projectes de aprenentatge automàtic.

Tot i això, hem pogut desenvolupar correctament un sistema d'aprenentatge per a la realització d'unes tasques específiques i hem pogut comprovar, encara que els resultats no fossin els ideals, que el funcionament era acceptable. S'ha intentat que la memòria plasmi la metodologia seguida i no es centri únicament en els resultats que estan lluny de poder substituir a interioristes per a generar moodboards.

Partint de la hipòtesi proposada al principi podem afirmar que:

- Hem pogut validar la interfície gràfica realitza molt més ràpid el procés de generació de moodboards, aquest procés ha passat de realitzar-se en 3 hores a realitzar-se en 1 hora.
- Hem pogut validar que començant des d'un moodboard generat d'una RBM, un expert pot editar els erros que hi troba i així anar més ràpid que si busques aquests ítems des de zero. Aquests moodboards serveixen com una plantilla bàsica per tal de augmentar la velocitat de creació dels moodboards.
- Podem utilitzar la màquina creada, per tal que donada una configuració fixa



d'elements en una graella, generar altres ítems sense modificar la distribució de la graella i així obtenir distribucions noves per tal de modificar-les en el futur.

Finalment, durant la realització del projecte han anat apareixent problemes que han provocat canvis importants en la implementació de la màquina i s'han anat canviat especificacions que afectaven a les dades. Finalment s'han definit uns estàndards per tal de poder fer funcionar correctament la màquina i obtenir resultats (encara que no fossin els ideals) per tal de poder comprovar que s'havia realitzat la implementació correctament.

Per acabar, el projecte ha suposat un repte tant per el seu desenvolupament com per la seva avaluació ja que en un principi es desconeixia com s'avaluaria el problema presentat ja que mai havia tractat un tema com aquest ni havia desenvolupat cap màquina d'aquest estil. A més la empresa no tenia cap idea clara de com avaluaria el projecte i per tant és difícil decidir si s'ha obtingut un resultat suficientment bo com per posar-ho en producció.

El projecte m'ha permès aprendre molt sobre el machine learning i la seva utilització. També m'ha forçat a aprendre noves eines vitals per a la implementació eficient del model, com la llibreria numpy i scipy que han sigut importants per a poder codificar les dades com a matrius sparse i poder entrenar el model de manera eficient.

### 6.3 Treball futur: Com millorar el projecte

Un cop s'ha acabat aquest projecte podem millorar-lo i seguir-lo desenvolupant de diverses maneres. En primer lloc i el següent pas que s'hauria de fer per a continuar el projecte seria que un cop generats els moodboards s'editessin per tal de corregir els objectes que són incorrectes i així obtenir nous moodboards per entrenar la màquina. D'aquesta forma s'obtindria un volum més elevat de dades que beneficiaria a la màquina ja que l'entrenament seria molt més correcte i les dades d'entrada més extenses.

Una altra opció seria generar les màquines per a diferents estils, ja que en aquest projecte només l'hem creat per un estil concret i poder seria interessant tenir tantes màquina com estils per tal de generar qualsevol estil de moodboard que es desitgi.

Una altra millora que es podria aplicar seria la de crear una RBM nova que a part de llegir els objectes també relacionés els objectes col·locats, és a dir, que establir normes mitjançant una xarxa neuronal per tal d'obtenir resultats més correctes. També es podria realitzar una ingesta per les imatges, és a dir, passar com a input les imatges en comptes de la informació dels objectes i així obtenir directament imatges de la sortida de la màquina. Una altra opció seria passar com a input la imatge i la graella per tal que l'aprenentatge tingues en compte les dues fonts d'informació i pogués avaluar millor les normes que s'estableixen..

Finalment una de les últimes millores que es podrien aplicar al projecte seria que tot estigués inclòs en l'aplicació, és a dir, que l'aplicació tingués ja implementades les màquines i que des de l'aplicació poguessin aprendre directament i generar els moodboards i mostrar-los. Això permetria un control sobre la visualització dels moodboards i sobre les dades d'entrenament ja que en tot moment es visualitzarien els objectes que es col·loquen i/o es generen.

## Referències

- [1] A Beginner's Tutorial for Restricted Boltzmann Machines,  
<https://deeplearning4j.org/restrictedboltzmanmachine>.
- [2] Eric Yuan; Restricted Boltzmann Machine,  
<http://eric-yuan.me/rbm/>, 2014.
- [3] Restricted Boltzmann Machines (RBM),  
<http://deeplearning.net/tutorial/rbm.html>.
- [4] Geoffrey Hinton; Advanced Machine Learning, Lecture 4, Restricted Boltzmann Machines,  
<https://www.cs.toronto.edu/~hinton/csc2535/notes/lec4new.pdf>.
- [5] Documentation RBM sklearn,  
[http://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.BernoulliRBM.html](http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.BernoulliRBM.html).
- [6] Documentation OneHotEncoder sklearn,  
<http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>.
- [7] Salakhutdinov, R. Learning Deep Generative Models.  
<https://www.cs.cmu.edu/~rsalakhu/papers/annrev.pdf>.
- [8] Representational power of restricted Boltzmann machines and deep belief networks  
<http://www.iro.umontreal.ca/~lisa/publications2/index.php/attachments/single/22>
- [9] Carreira-Perpiñán, M. A. and Hinton, G. E. On Contrastive Divergence Learning  
<http://www.cs.toronto.edu/~fritz/absps/cdmiguel.pdf>
- [10] Ranzato, M., Krizhevsky, A., and Hinton, G. E.  
[http://www.cs.toronto.edu/~fritz/absps/ranzato\\_aistats2010.pdf](http://www.cs.toronto.edu/~fritz/absps/ranzato_aistats2010.pdf)