

Un sistema de recomendación basado en factorización de matrices y descenso en gradiente estocástico

Mario Becerra Contreras

Objetivo

Implementar un sistema de recomendación utilizando factorización de matrices utilizando como método de optimización descenso en gradiente estocástico.

Sistemas de recomendación

Pretenden predecir las respuestas de los usuarios a distintas opciones. Dos tipos:

1. Basados en contenido: A partir de características y propiedades de los productos (por ejemplo, género, actores, país de origen, año, etc.).
2. Colaborativos: A partir de usuarios y productos se construyen medidas de similitud en el sentido de que les han gustado los mismos productos o que les gustaron a las mismas personas.

Ejemplos de sistemas de recomendación

- Ofrecer películas que a un usuario le pueden gustar
- Sugerir artículos que un usuario probablemente compre
- Recomendar canciones en un servicio de *streaming*

Datos

Calificaciones explícitas por usuario en matriz R

	P_1	P_2	P_3	\dots	P_n
U_1	1	2	3	\dots	-
U_2	4	-	4	\dots	-
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
U_m	-	-	1	\dots	3

Objetivo: llenar espacios faltantes

Colaborativos: modelo base

Supone que artículos y usuarios tienen sesgos.

Buena parte de la calificación observada se debe a efectos asociados a los sesgos.

Predicción de calificación de usuario i a artículo j :

$$\hat{r}_{ij} = \mu + a_i + b_j$$

M_i = Número de artículos calificados por usuario i

N_j = Número de calificaciones de artículo j

$$a_i = \frac{1}{M_i} \sum_t r_{it} - \mu$$

$$b_j = \frac{1}{N_j} \sum_s r_{sj} - \mu$$

Colaborativos: factorización de matrices

Idea: Suponer factores ocultos o latentes que caracterizan usuarios y productos.

Por ejemplo, factor de comedia-seriedad y factor de fantasía-realidad.

Usuarios tienen valores en cada factor latente dependiendo de la inclinación que tengan; artículos son representados por qué tanto tienen de cada factor latente.

La calificación un artículo es una combinación lineal de factores de usuarios y artículos.

Colaborativos: factorización de matrices

Representación matricial de factores latentes en usuarios (U) y artículos (P).

Supongamos dos factores latentes: seriedad-comedia y fantasía-realidad.

Al usuario 1 le gustan las películas serias de fantasía y al 2 le gustan las comedias pero es indiferente entre fantasía o realidad.

The diagram illustrates the factorization of user preferences into latent factors. It shows two matrices, P and U , each represented as a product of a vector and a matrix of latent factors.

Matrix P (Articles):

$$P = \begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} = \begin{bmatrix} \boxed{2} & \boxed{4} \\ \boxed{-3} & \boxed{-3} \end{bmatrix}$$

Annotations for P :

- Red arrow from $\boxed{2}$ to *El Señor de los Anillos*
- Red arrow from $\boxed{-3}$ to *Mean girls*
- Blue arrow from the first column ($\boxed{2}, \boxed{-3}$) to *Seriedad-Comedia*
- Blue arrow from the second column ($\boxed{4}, \boxed{-3}$) to *Fantasía-Realidad*

Matrix U (Users):

$$U = \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} = \begin{bmatrix} \boxed{3} & \boxed{4} \\ \boxed{-2} & \boxed{0} \end{bmatrix}$$

Annotations for U :

- Blue arrow from the first column ($\boxed{3}, \boxed{-2}$) to *Seriedad-Comedia*
- Blue arrow from the second column ($\boxed{4}, \boxed{0}$) to *Fantasía-Realidad*

Colaborativos: factorización de matrices

Calificación resultante de combinación lineal:

$$R = \begin{bmatrix} u_1^T p_1 & u_1^T p_2 \\ u_2^T p_1 & u_2^T p_2 \end{bmatrix} = \begin{bmatrix} 22 & -21 \\ -4 & 6 \end{bmatrix}$$

Colaborativos: factorización de matrices

En la vida real no se tienen las matrices U y P , sino la matriz R . Las matrices U y P se estiman a partir de la matriz R .

Las matrices U y P son de rango bajo tales que

$$R \approx UP^T$$

Función de pérdida:

Predicción:

$$r_{ij} = u_i^T p_j$$

$$RMSE = \left(\frac{1}{|A|} \sum_{(i,j) \in A} (r_{ij} - u_i^T p_j)^2 \right)^{\frac{1}{2}}$$

A es el conjunto de usuarios para los cuales se conoce la calificación

Colaborativos: factorización de matrices con sesgos

Agregar idea de modelo base: existen sesgos en los artículos y en los usuarios

Predicción:

$$r_{ij} = \mu + a_i + b_j + u_i^T p_j$$

Problema a resolver:

$$\min_{p,u,b,a} \sum_{(i,j) \in A} \left[(r_{ij} - \mu - a_i - b_j - u_i^T p_j)^2 + \lambda (\|p_j\|^2 + \|u_i\|^2 + a_i^2 + b_j^2) \right]$$

Colaborativos: factorización de matrices con sesgos

$$\begin{aligned} L &= \sum_{(i,j) \in A} \left[(r_{ij} - \mu - a_i - b_j - u_i^T p_j)^2 + \lambda \left(\|p_j\|^2 + \|u_i\|^2 + a_i^2 + b_j^2 \right) \right] \\ &= \sum_{(i,j) \in A} \left[(x_{ij} - a_i - b_j - u_i^T p_j)^2 + \lambda \left(\|p_j\|^2 + \|u_i\|^2 + a_i^2 + b_j^2 \right) \right] \\ &= \sum_{(i,j) \in A} (x_{ij} - a_i - b_j - u_i^T p_j)^2 \\ &\quad + \lambda \left(\sum_j \sum_{m=1}^k p_{jm}^2 + \sum_i \sum_{m=1}^k u_{im}^2 + \sum_i a_i^2 + \sum_j b_j^2 \right). \end{aligned}$$

Descenso en gradiente estocástico

$$\ell_{ij} = (x_{ij} - a_i - b_j - u_i^T p_j)^2 + \lambda \left(\sum_{m=1}^k p_{jm}^2 + \sum_{m=1}^k u_{im}^2 + a_i^2 + b_j^2 \right)$$

$$L = \sum_{(i,j) \in A} \ell_{ij}$$

Descenso en gradiente estocástico

$$\frac{\partial \ell_{ij}}{\partial u_{im}} = -2e_{ij}p_{jm} + 2\lambda u_{im}$$

$$\frac{\partial \ell_{ij}}{\partial p_{jm}} = -2e_{ij}u_{im} + 2\lambda v_{jm}$$

$$\frac{\partial \ell_{ij}}{\partial b_j} = -2e_{ij} + 2\lambda b_j$$

$$\frac{\partial \ell_{ij}}{\partial a_i} = -2e_{ij} + 2\lambda a_i$$

Donde:

$$e_{ij} = x_{ij} - a_i - b_j - u_i^T p_j$$

Descenso en gradiente estocástico

Iniciar vectores a , b y matrices P y Q

mientras *no se cumpla el criterio de paro* **hacer**

para todo $(i, j) \in A$ **hacer**

para todo $m \in 1, \dots, k$ **hacer**

$u_{im} \leftarrow u_{im} - \gamma (-2e_{ij}p_{jm} + 2\lambda u_{im})$

$p_{jm} \leftarrow p_{jm} - \gamma (-2e_{ij}u_{im} + 2\lambda v_{jm})$

fin

$a_i \leftarrow a_i - \gamma (-2e_{ij} + 2\lambda a_i)$

$b_j \leftarrow b_j - \gamma (-2e_{ij} + 2\lambda b_j)$

fin

fin

Evaluación de sistema

Seleccionar cada elemento i calificado con la calificación máxima por cada usuario u en el conjunto de prueba:

1. Seleccionar 1000 artículos adicionales que no fueron calificados por el usuario u .
2. Calcular la calificación de acuerdo al modelo para el artículo i y cada uno de los 1000 artículos adicionales.
3. Formar lista ordenada de acuerdo a las calificaciones predichas por el modelo para los 1001 artículos. Sea p el lugar que ocupa el artículo de prueba i entre todos los elementos de la lista. El mejor resultado corresponde al caso en que el artículo de prueba i está antes de todos los artículos aleatorios, i.e., $p = 1$.
4. Se forma una lista de las mejores N recomendaciones al seleccionar los N artículos mejor calificados en la lista. Si $p \leq N$ entonces se tiene un acierto, en otro caso se tiene un desacierto. La probabilidad de obtener aciertos aumenta si N aumenta, y si $N = 1001$, siempre se tiene un acierto.

Evaluación de sistema

$$\text{recall}(N) = \frac{\#hits}{|T|}$$

$$\text{prec}(N) = \frac{\#hits}{N \cdot |T|} = \frac{\text{recall}(N)}{N}$$

Resultados: *MovieLens*

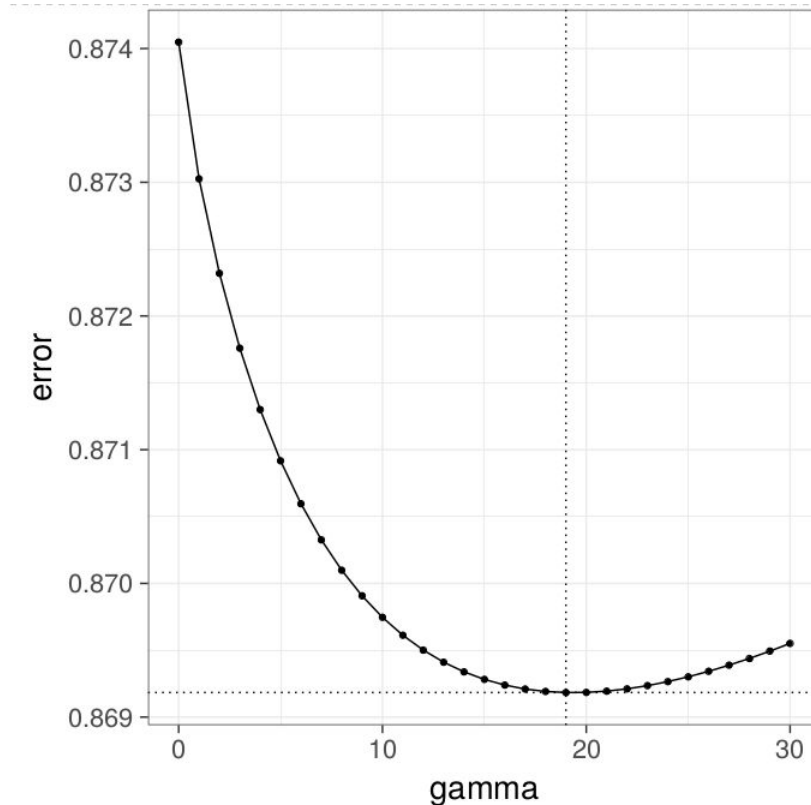
20,000,263 calificaciones de 26,744 películas hechas por 138,493 usuarios

Conjunto	Artículos	Usuarios	Calificaciones
Entrenamiento	26,247	138,493	18,029,206
Validación	6,256	41,483	1,469,158
Prueba	2,895	20,676	501,899

Resultados: *MovieLens*

Modelo base:

error en conjunto de prueba: 0.8755

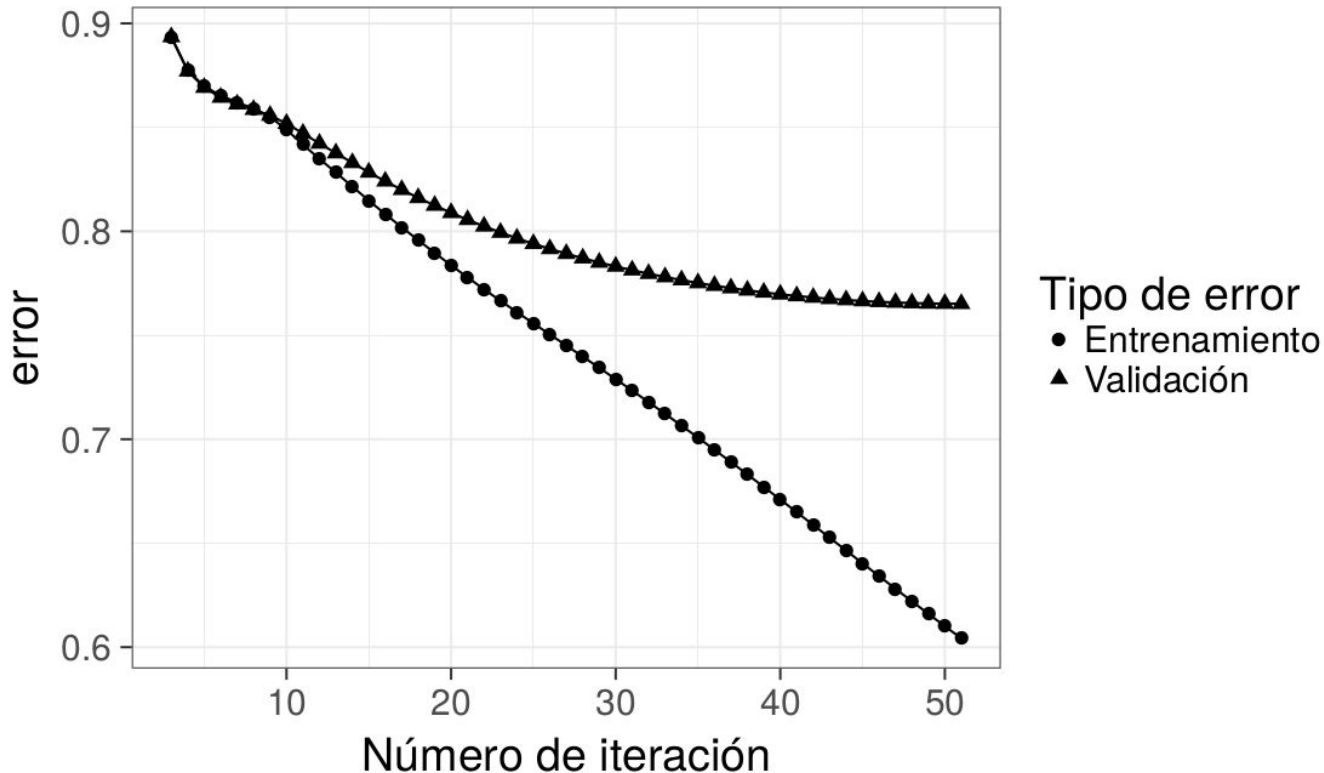


Resultados: *MovieLens*

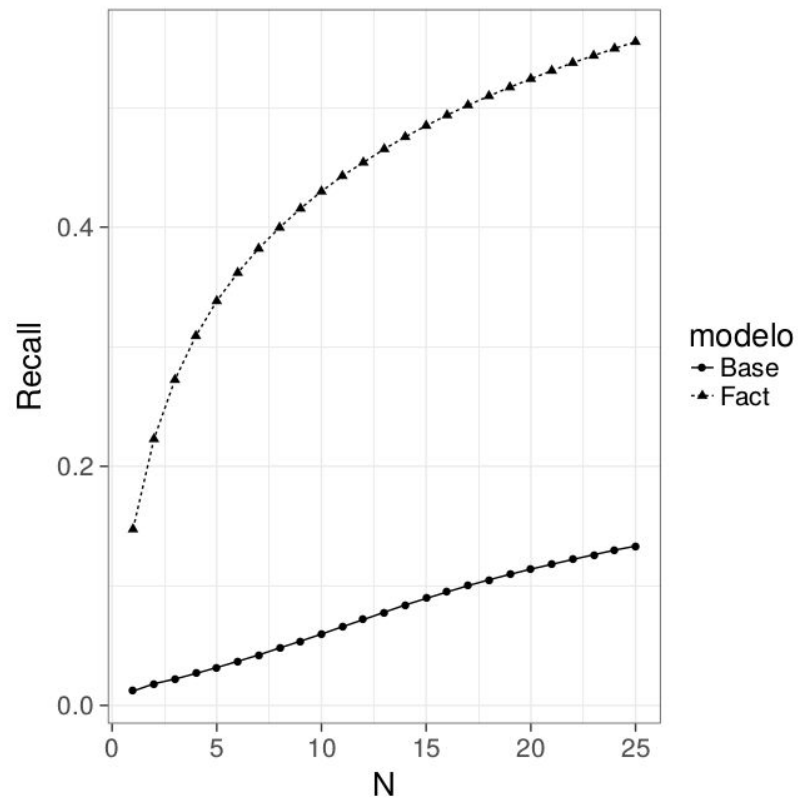
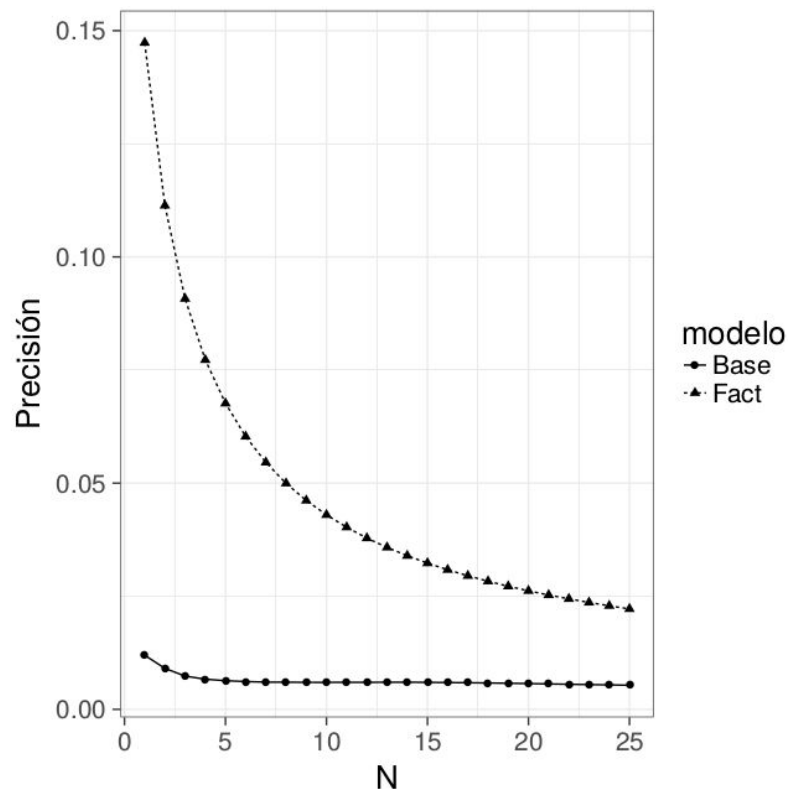
Modelo final:

200 dimensiones latentes,
tasa de aprendizaje de
0.001 y $\lambda = 0.01$

Error en conjunto de
prueba de 0.7651



Resultados: *MovieLens*



Resultados: *MovieLens*

Aladdin	Lion King, The	1.74
	Beauty and the Beast	1.79
	Little Mermaid, The	2.09
	The Princess and the Frog	2.33
	Tarzan	2.42
	Mulan	2.43
	Oliver & Company	2.47
	Great Mouse Detective, The	2.47
	How to Train Your Dragon 2	2.49
	Anastasia	2.49

Resultados: *MovieLens*

Harry Potter and the Sorcerer's Stone	HP and the Chamber of Secrets	0.47
	HP and the Prisoner of Azkaban	1.11
	HP and the Goblet of Fire	1.13
	HP and the Order of the Phoenix	1.33
	HP and the Half-Blood Prince	1.59
	HP and the Deathly Hallows: Part 2	1.83
	HP and the Deathly Hallows: Part 1	1.87
	Narnia: The Voyage of ...	2.76
	Narnia: Prince Caspian	2.88
	The Lion King 1 1/2	2.91

Resultados: *MovieLens*

Pulp Fiction	Reservoir Dogs	2.26
	Inglorious Bastards	3.03
	Goodfellas	3.15
	An Evening with Kevin Smith 2	3.19
	Black Mirror	3.2
	Ricky Gervais Live 3: Fame	3.23
	Nightcrawler	3.24
	Inside Llewyn Davis	3.25
	Flirting	3.25
	Generation Kill	3.25

Resultados: *BookCrossing*

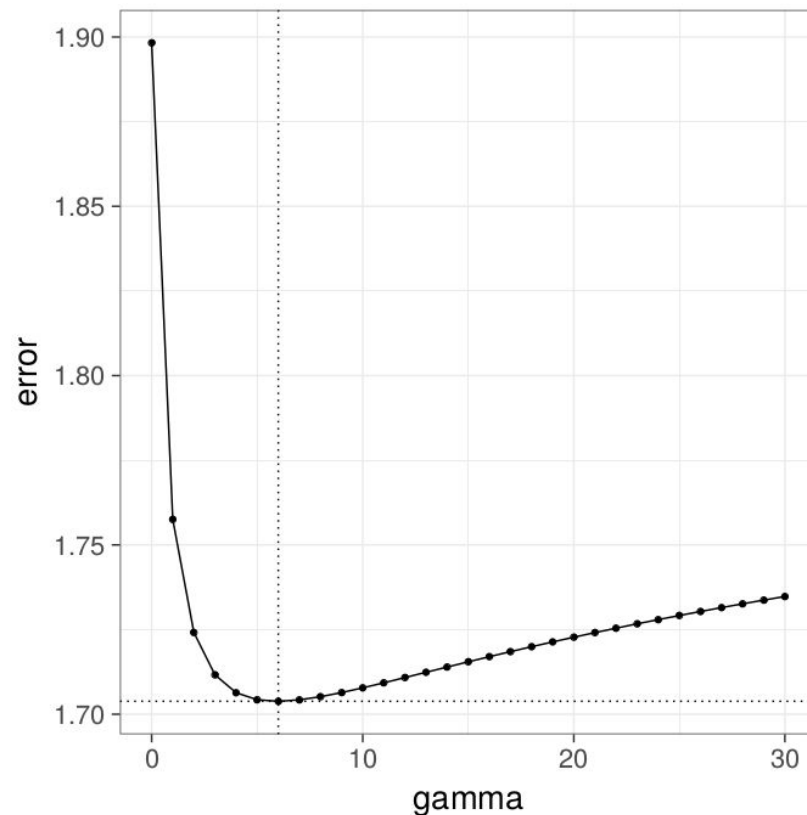
383,852 calificaciones de 153,683 libros hechas por 68,092 usuarios

Conjunto	Artículos	Usuarios	Calificaciones
Entrenamiento	140,807	64,459	351,217
Validación	13,072	5,332	21,224
Prueba	5,065	2,286	7,435

Resultados: *BookCrossing*

Modelo base:

error en conjunto de prueba: 1.6977

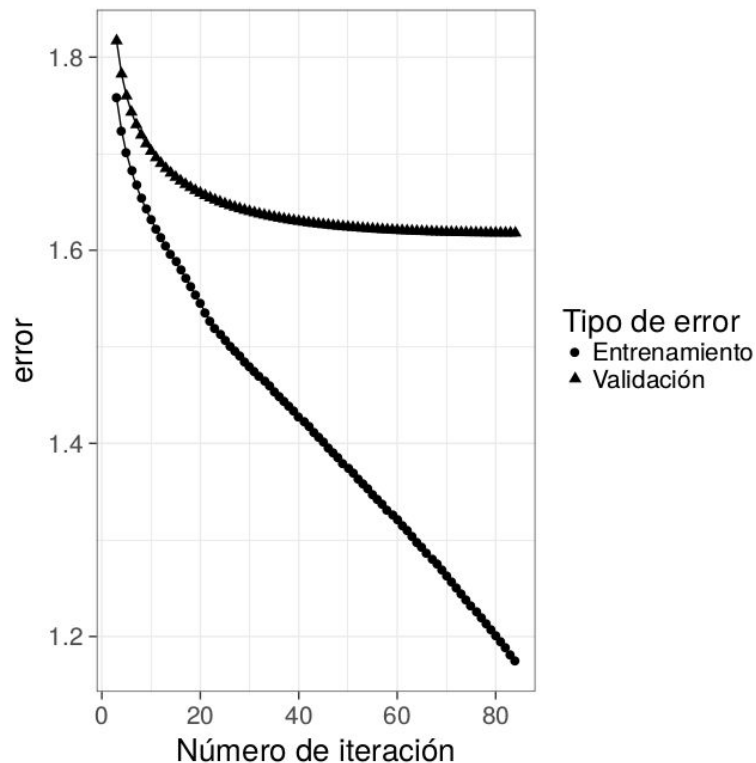


Resultados: *MovieLens*

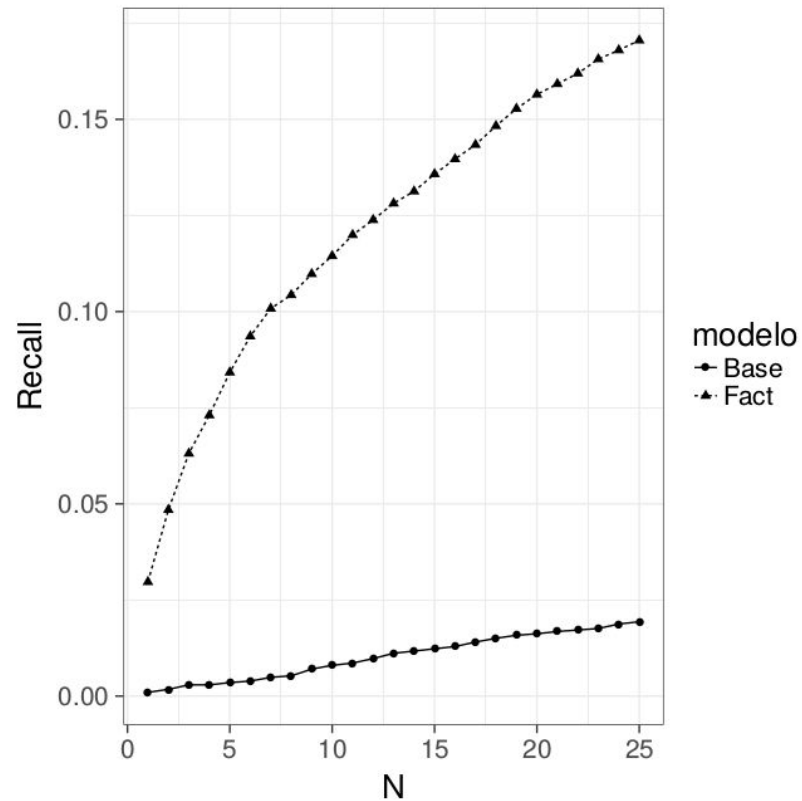
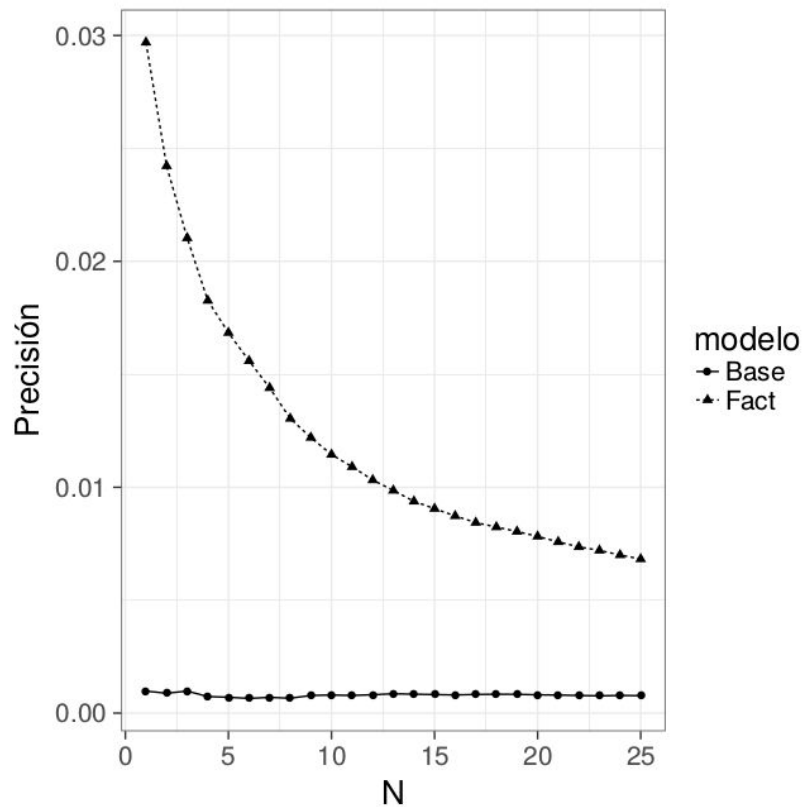
Modelo final:

200 dimensiones latentes,
tasa de aprendizaje de
0.001 y $\lambda = 0.01$

Error en conjunto de
prueba de 1.5865



Resultados: *BookCrossing*



Resultados: *BookCrossing*

Harry Potter and the Sorcerer's Stone	Who in Hell Is Wanda Fuca?	1.71
	A Ranking of the Most Influential Persons...	1.75
	Harry Potter and the Goblet of Fire (Book 4)	1.77
	Surviving Production: The Art of Production ...	1.77
	Just Tell Me When We're Dead	1.79
	The Conquest	1.8
	The Ghosts of Cougar Island	1.8
	Dr. Death (Alex Delaware Novels (Paperback))	1.81
	Familiar Obsession (Fear Familiar)	1.81
	My Gal Sunday	1.81

Resultados: *BookCrossing*

The Fellowship of the Ring	I Hate Texas: 303 Reasons Why You...	0.69
	King Lear	0.73
	When Your Marriage Needs Repair	0.73
	365 Things Every Couple Should Know	0.73
	So You Want to Be a Witch	0.74
	Keeping of Customers: A Treasury of ...	0.74
	Zin! Zin! Zin! A Violin	0.74
	Midsummer Nights Dream	0.75
	Pinky And Rex And The Mean Old Witch	0.75
	Miss Nelson Is Missing!	0.75

¡Gracias!