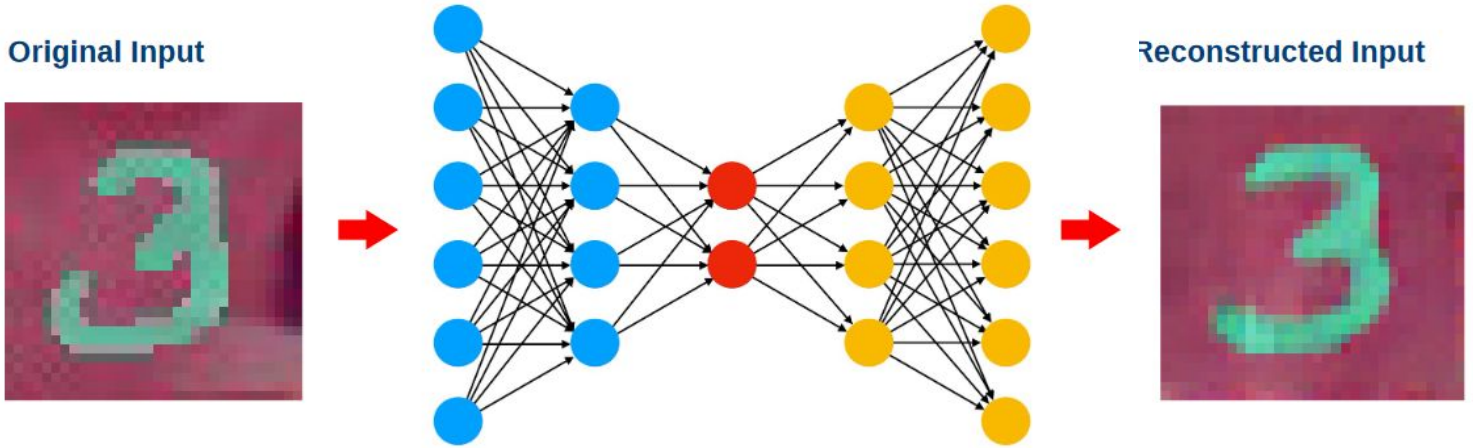
Easer

Michiel téblick
Thibaut Van Goethem

Introduction

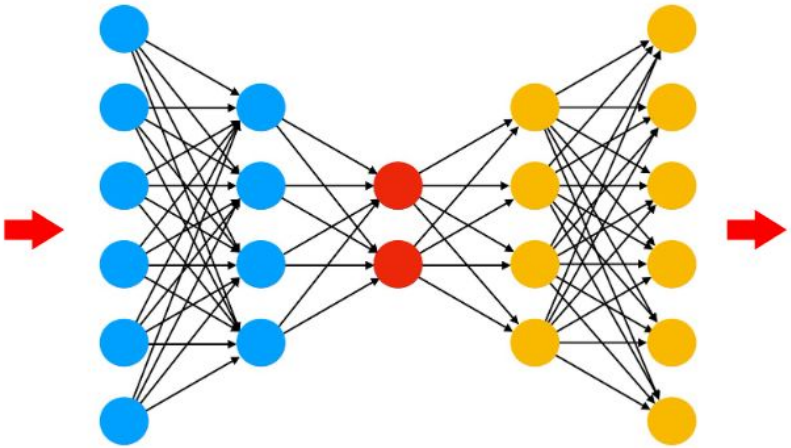
- Autoencoder



Introduction

- **Autoencoder**

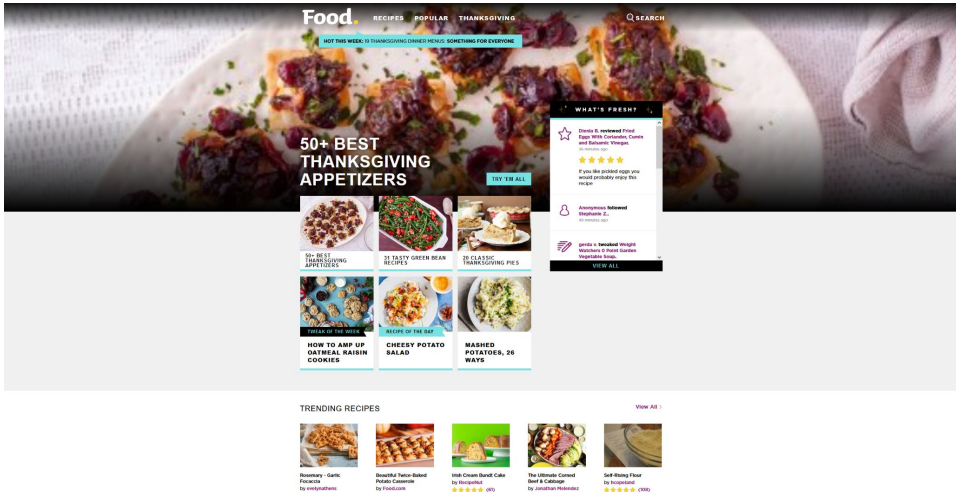
user interactions



user recommendations

Introduction

- **dataset**
  - food.com
  - 1132367 interactions
  - 231637 recipes



Formula

- **prediction**
  - rec=interactions*B
    - B -> [|items|,|items|]
    - interactions -> [|rec users|,|items|]
    - rec -> [|rec users|,|items|]

Formula

- **prediction**
  - rec=interactions*B
    - B -> [|items|,|items|]
    - interactions -> [|rec users|,|items|]
    - rec -> [|rec users|,|items|]

$$\begin{array}{ccc} a & b & c \\ (0 & 1 & 0) \end{array} * \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} = \begin{array}{ccc} a & b & c \\ (1 & 0 & 0) \end{array}$$

Formula

- learning B

Formula

- **learning B**
  - $$\min_{B} \quad ||X - XB||_F^2 + \lambda \cdot ||B||_F^2$$
    $$\text{s.t.} \qquad \text{diag}(B) = 0$$
    - B -> [|items|,|items|]
    - X -> [|users|,|items|]
    - λ -> single value (regularization)

Formula

- learning B
  - $$\min_{B} \quad ||X - XB||_F^2 + \lambda \cdot ||B||_F^2$$
    $$\text{s.t.} \qquad \text{diag}(B) = 0$$
  - $$\hat{P} \triangleq (X^\top X + \lambda I)^{-1}$$
  - $$\hat{B}_{i,j} = \begin{cases} 0 & \text{if } i = j \\ -\frac{\hat{P}_{ij}}{\hat{P}_{jj}} & \text{otherwise.} \end{cases}$$

Interpretation

$$\hat{B}_{i,j} = \begin{cases} 0 & \text{if } i = j \\ -\frac{\hat{P}_{ij}}{\hat{P}_{jj}} & \text{otherwise.} \end{cases}$$

- **Interpretation of P**

$$\hat{P} \triangleq (X^\top X + \lambda I)^{-1}$$

Drop L2-norm regularization

$$\hat{P} \approx (X^T X)^{-1}$$

$X^TX \approx$ covariance matrix Σ

$$\hat{P} \approx \Sigma^{-1}$$

Interpretation

$$\hat{B}_{i,j} = \begin{cases} 0 & \text{if } i = j \\ -\frac{\hat{P}_{ij}}{\hat{P}_{jj}} & \text{otherwise.} \end{cases}$$

- **Interpretation of XB**

$$\begin{aligned} E[x_j | x_{-j}] &= -x_{-j} \cdot P_{-j,j} / P_{j,j} \\ &= x_{-j} \cdot B_{-j,j} \\ &= x \cdot B_{.,j} \end{aligned}$$

$$x_j \cdot B_{j,j} = 0$$

Results paper

- **MovieLens data (136,667 users, 20,108 movies and 10 million interactions)**

	Recall@20	Recall@50	NDCG@100
Popularity	0.162	0.235	0.191
$EASE^R$	0.391	0.521	0.420
SLIM	0.370	0.495	0.401
WMF	0.360	0.498	0.386
CDAE	0.391	0.523	0.418
$MULT\text{-}VAE^{PR}$	0.395	0.537	0.419

Results paper

- **Netflix data (463,435 users, 17,569 movies and 57 million interactions)**

	Recall@20	Recall@50	NDCG@100
Popularity	0.116	0.175	0.159
$EASE^R$	0.362	0.445	0.393
SLIM	0.347	0.428	0.379
WMF	0.316	0.404	0.351
CDAE	0.343	0.428	0.376
$MULT\text{-}VAE^{PR}$	0.351	0.444	0.386

Results paper

- **Million Song data (571,355 users, 41,140 songs and 34 million interactions)**

	Recall@20	Recall@50	NDCG@100
Popularity	0.043	0.068	0.058
$EASE^R$	0.333	0.428	0.389
SLIM	did not finish	did not finish	did not finish
WMF	0.211	0.312	0.257
CDAE	0.188	0.283	0.237
$MULT\text{-}VAE^{PR}$	0.266	0.364	0.316

Our results

- **Pre-processing recipes dataset**

226570 users, 231637 recipes and 1 million interactions

22729 users, 16609 recipes and 404811 interactions

Our results

	Recall@20	Recall@50	ndcg@100
Popularity	0.001	0.003	0.001
$EASE^R$	0.044	0.077	0.031

Demonstration