Recommendation Systems

Recommendation systems are everywhere

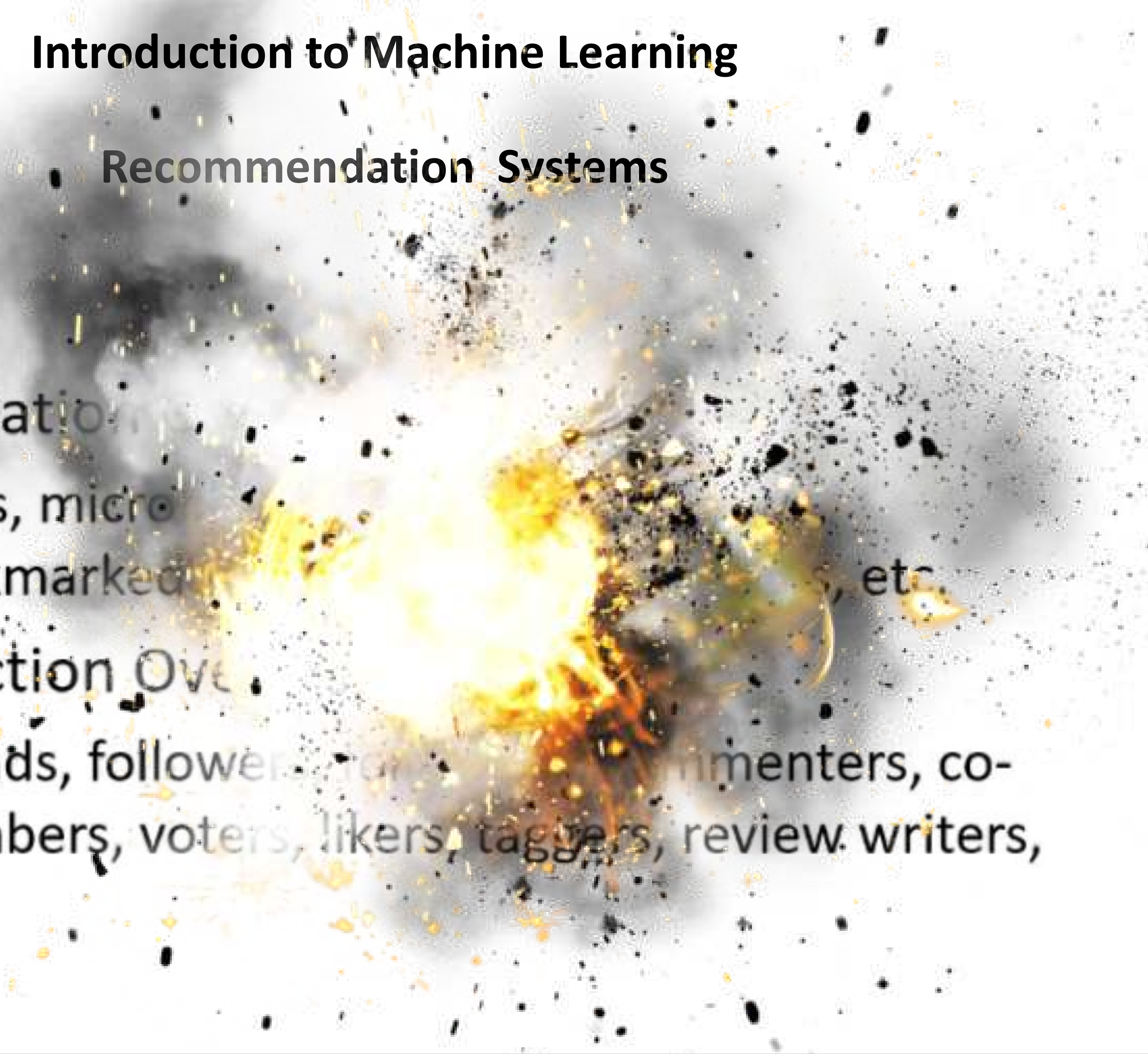
Recommendation Systems

Social overload

- Facebook ... est socia... ork site
  - 600,00...
  - 35,000,00...
  - 900,000,0...
  - 30,000,000 ... nth
- YouTube – lar...
  - 2,000,000,000...
  - 1,000,000 vide... month
- Twitter – largest microblogging site
  - 200,000,000 users per month
  - 65,00,000 tweets per day (750 per second)
  - 8,000,000 followers of most popular user

Recommendation Systems

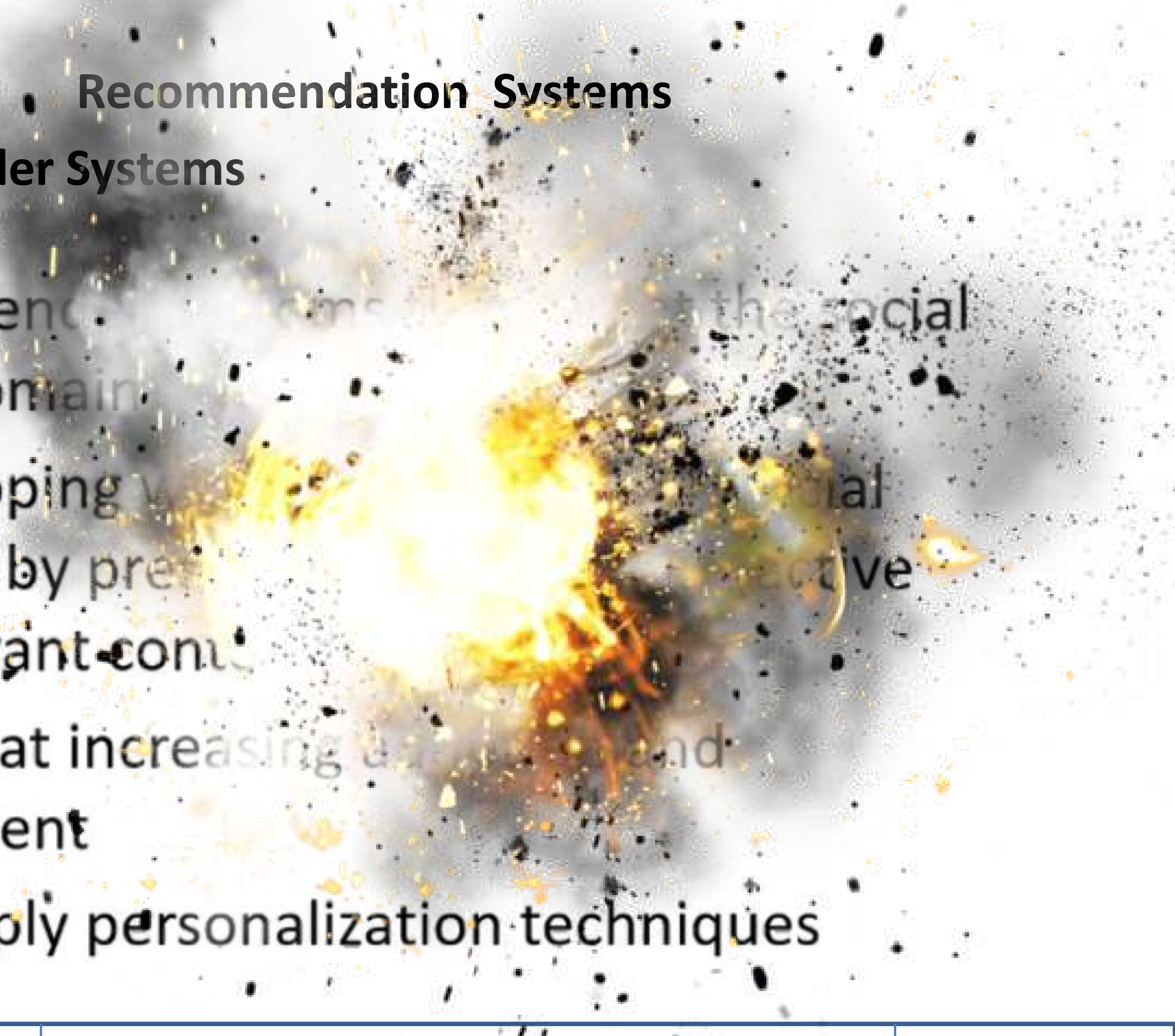
Social overload



- Informati
  - Blogs, micr
    bookmark … , etc.
- Interaction Ov
  - Friends, followe … menters, co-members, vote …, likers, tagg … s, review writers, etc.

Recommendation Systems

Social Recommender Systems

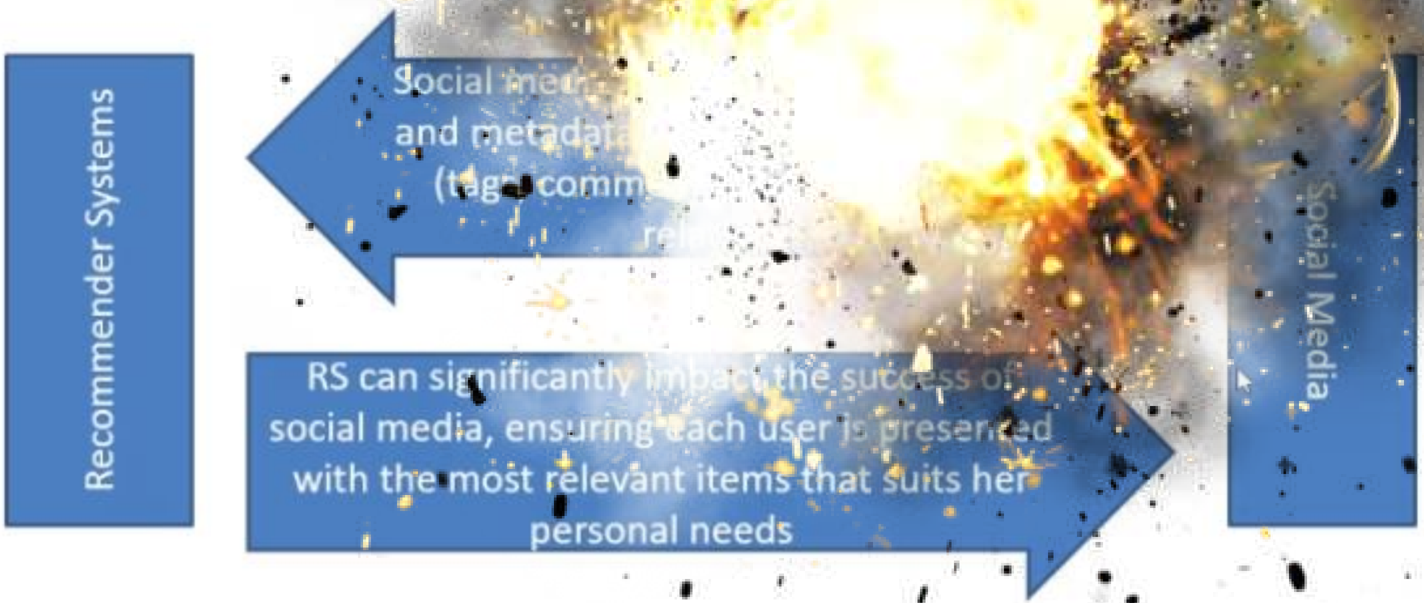


- Recommen[illegible] the social media domain
- Aim at coping w[illegible]al overload by pre[illegible]ctive and relevant cont[illegible]
- Also aim at increasing [illegible] and engagement
- Often apply personalization techniques

Recommendation Systems

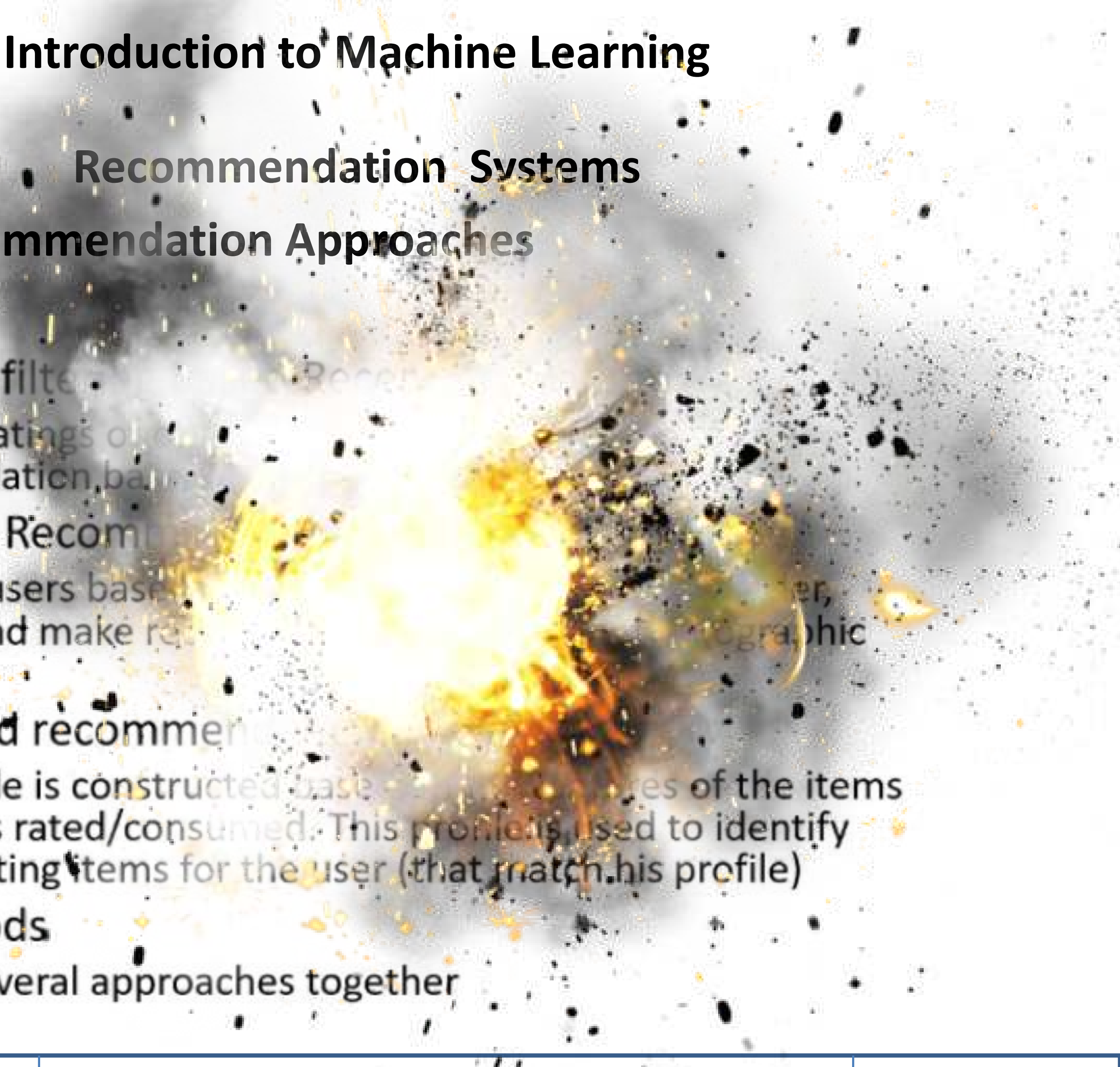
Recommender Systems & Social Media

- Recommender Systems ... augmentation ... social process in which we rely on a... suggestions ... people
- Social Media and Reco... other



Recommendation Systems

Fundamental Recommendation Approaches



- Collaborative filte
  - Aggregate ratings o
    recommendation ba
- Demographic Recom
  - Categorize users bas er,
    income..) and make re ographic
    classes
- Content-based recommen
  - A user profile is construct es of the items
    the user has rated/consumed. This profile is used to identify
    new interesting items for the user (that match his profile)
- Hybrid methods
  - Combine several approaches together

Recommendation Systems

Collaborative Filtering

- In the real world we seek adv[illegible] [illegible]ends, colleagues, experts)
- CF automates the process of "w[illegible]
  - Weight all users with respect to similarity with the active user.
  - Select a subset of the users (neighbors) to use as recommenders
  - Predict the rating of the active user for specific items based on its neighbors' ratings
  - Recommend items with maximum prediction

Recommendation Systems

User based Collaborative Filtering Algorithm

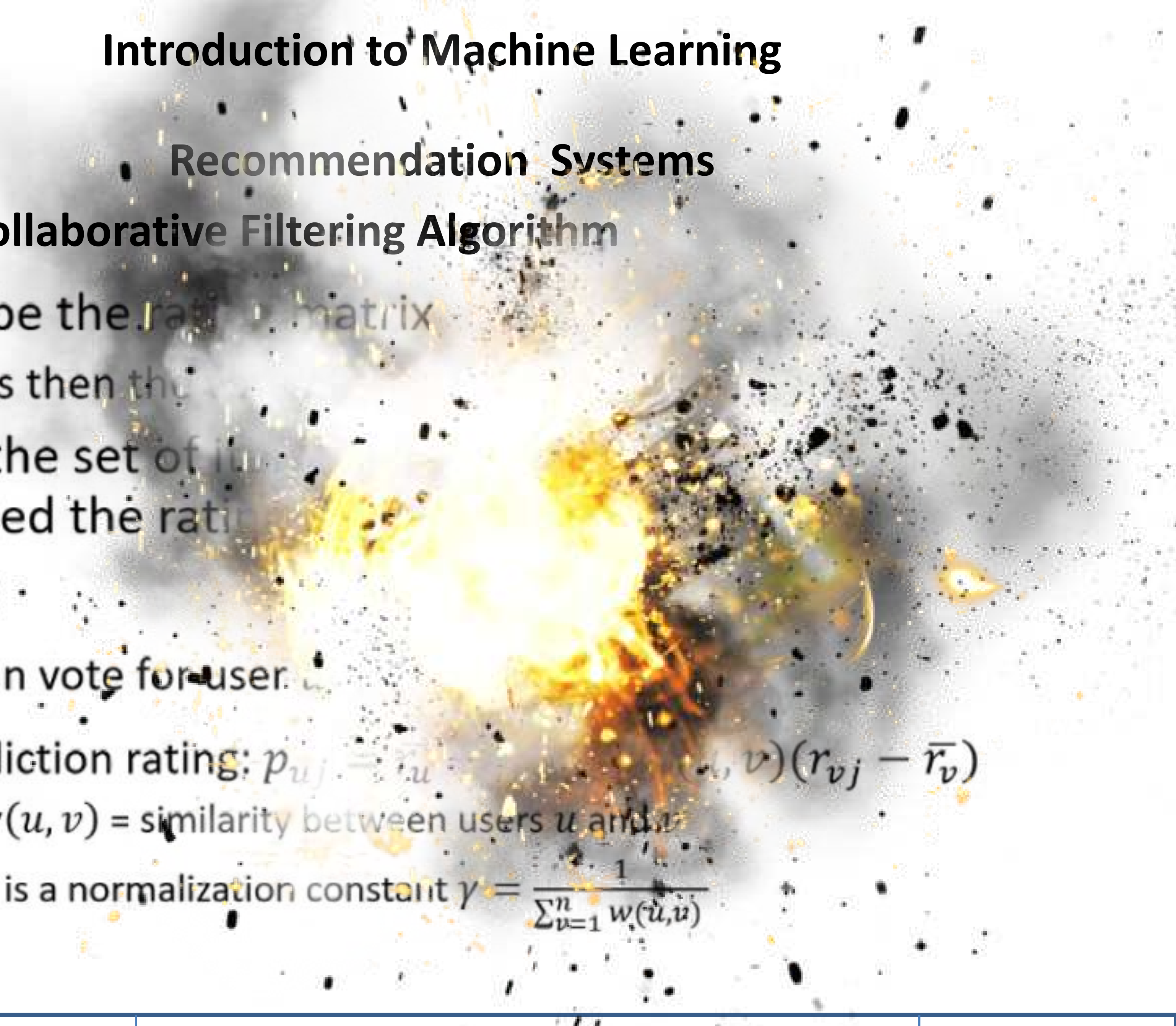
- The User x It... rix

	Shr...
Alice	Like
Bob	?
Chris	Like
John	Like

- Shall we recommen... ohn?
- John's taste is sim... to bo... and Alice tastes ⇒ Do not recommend Superman to John

Recommendation Systems

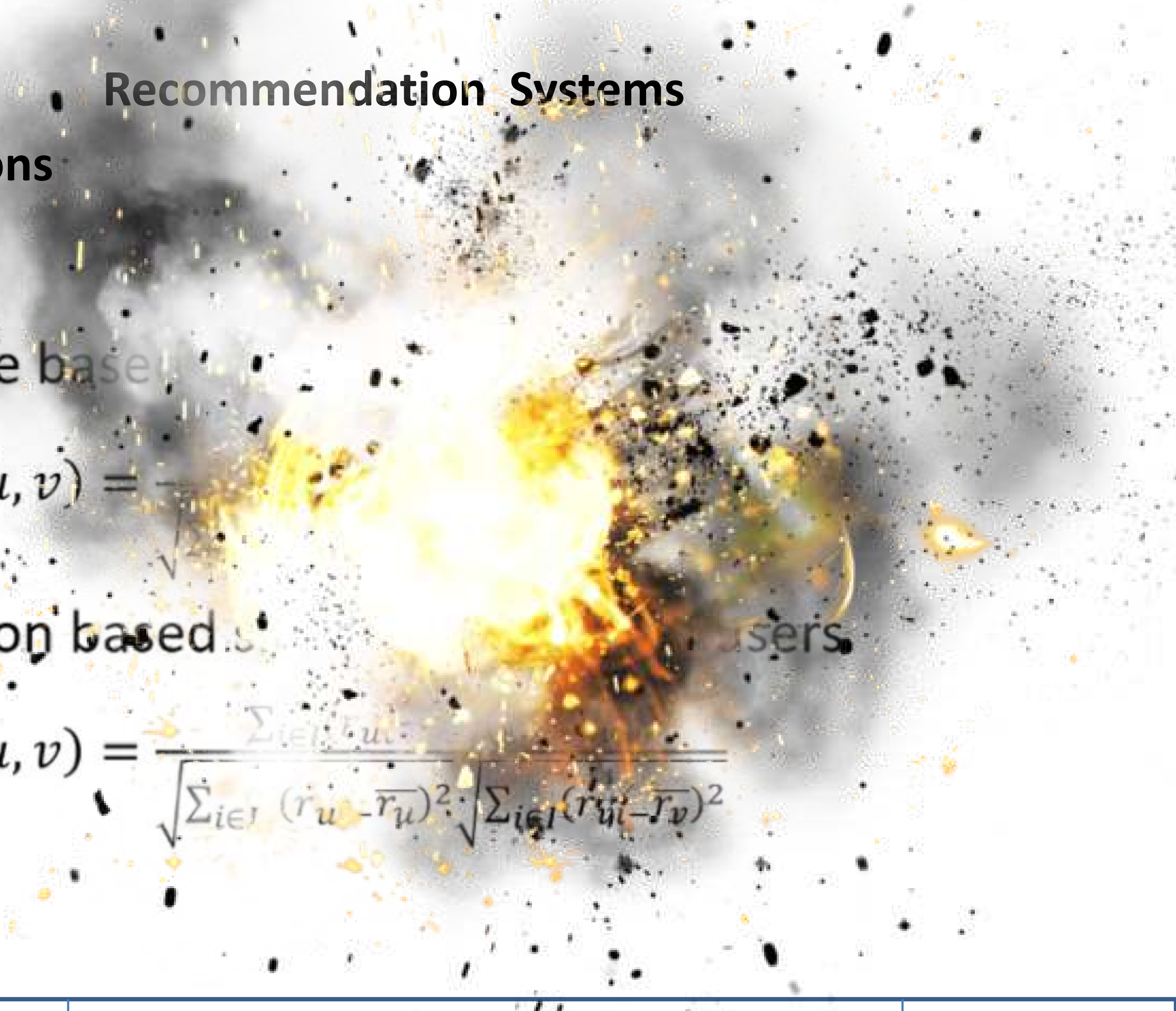
User based Collaborative Filtering Algorithm



- Let R be the ra… matrix
  - r_{uj} is then th…
- I_u be the set of i… provided the rati…
- Voting
  - Mean vote for user …
  - Prediction rating: $p_{uj} = \bar{r}_u$ … $(u, v)(r_{vj} - \bar{r}_v)$
    - $w(u, v)$ = similarity between users u and v
    - γ is a normalization constant $\gamma = \frac{1}{\sum_{v=1}^{n} w(u,v)}$

Recommendation Systems

Similarity Functions



- Cosine bas…
  - $w(u, v) = \dots$
- Pearson based s… users
  - $w(u, v) = \frac{\sum_{i \in I} r_{u} \dots}{\sqrt{\sum_{i \in I}(r_{u} - \overline{r_u})^2}\sqrt{\sum_{i \in I}(r_{vi} - \overline{r_v})^2}}$

Collaborative Filtering: Practical Challenges

- Ratings data is o[illegible]rse, an[illegible]of users with few co-ratings are pr[illegible]
- Fails to incorporat[illegible] the population as a wh[illegible]
  - Agreement abou[illegible]ss important than agr[illegible]
    - Some algorithms acc[illegible] by including weights inversely propo[illegible]ity.
- Calculating a user's per[illegible] is expensive – requiring comparison against all other users
  - Sampling: a subset of users is selected prior to prediction computation
  - Clustering: can be used to quickly locate a user's neighbors

Recommendation Systems

Item-Based Nearest Neighbor Algorithm

- The transpose of the user-based algorithms
  - Generate predictions [illegible] similarities betw[illegible]
  - The prediction for a[illegible] on the u[illegible] similar ite[illegible]

	[illegible]	[illegible]	[illegible]
Alice	[illegible]	[illegible]	[illegible]
Bob	[illegible]	[illegible]	[illegible]
Chris	[illegible]	[illegible]	[illegible]
John	Like	[illegible]	[illegible]

- Bob dislikes Snow-white (which i[illegible] recommend Shrek to Bob
- Predicted rating: $p_{uj} = \gamma \sum_{i=1}^{m} w(i,j) r_{ui}$
- Traverse over all m items rated by user u and measure their rating, averaged by their similarity to the predicted item
- w(i,j) is a measure of item similarity - usually the cosine measure
- Average correction is not needed because the component ratings are all from the same target user