



Universiteit
Leiden
The Netherlands

Information diffusion analysis in online social networks based on deep representation learning

Chen, X.

Citation

Chen, X. (2022, October 25). *Information diffusion analysis in online social networks based on deep representation learning*. Retrieved from <https://hdl.handle.net/1887/3484562>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3484562>

Note: To cite this publication please use the final published version (if applicable).

Nederlandse Samenvatting

Met de opkomst van online sociale netwerken (OSNs) is de manier waarop mensen informatie delen veranderd, wat sneller en breder wordt dan traditionele media. Begrijpen hoe informatie (zowel goede als schadelijke) zich via OSNs verspreidt, en welke elementen het succes van informatieverspreiding stimuleren, heeft belangrijke implicaties voor een breed scala aan toepassingen in de echte wereld. In dit proefschrift doen we onderzoek om de diffusie van informatie in OSNs te leren kennen door gebruik te maken van Deep Representation Learning. Concreet willen we twee specifieke taken oplossen, namelijk het modelleren van informatiecascades en het detecteren van geruchten.

We dragen het volgende bij:

We introduceren recurrent cascades convolutional network (CasCN) in Hoofdstuk 4, een op grafieken gebaseerd neuraal model voor voorspelling van informatiecascades op macroniveau. CasCN gebruikt een combinatie van een graafconvolutienetwerk en een terugkerend neuraal netwerk om de incrementele grootte van een cascade te voorspellen door structurele en tijdelijke kenmerken te extraheren uit een reeks op tijdstempels gebaseerde subgrafieken. CasCN introduceert ook CasLaplace voor gerichte grafieken, waarmee de beperkingen van eerdere neurale netwerken voor graafgrafieken worden overwonnen bij het omgaan met gerichte grafieken. De experimentele resultaten die zijn uitgevoerd op twee real-world datasets, laten zien dat CasCN zeer geschikt is voor het modelleren van structurele-tijdelijke kenmerken van cascades.

We introduceren multi-scale cascades model (MUCas) in Hoofdstuk 5, die gericht is op het vastleggen van multi-scale functies voor voorspelling van informatiecascades op macroniveau. MUCas gebruikt multi-scale graph capsule network en influence attention om de multi-scale informatie te leren en samen te smelten om een unieke cascade representatie. MUCas verbetert ook de bemonsteringsmethoden in CasCN door subgrafieken te bemonsteren op basis van tijdsintervallen in plaats van tijdstempels. We voeren experimenten uit met real-world datasets en tonen aan dat MUCas bijzonder effectief is in het extraheren van kenmerken op cascades van verschillende schalen, en multi-scale functies zijn van vitaal belang voor het verbeteren van de nauwkeurigheid van voorspellingen.

We introduceren macroscopic and microscopic-aware rumor detection model (MMRD) in Hoofdstuk 6, een op diffusie gebaseerd geruchtendetectie model dat geruchten detecteert door alleen verschillende niveaus van diffusie patronen te onderzoeken. MMRD maakt gebruik van graph neural networks om de macroscopische diffusie van geruchten propagatie te leren en microscopische diffusie patronen vast te leggen met behulp van bidirectional recurrent neural networks, rekening houdend met de user-time serie. MMRD maakt ook gebruik van de knowledge distillation techniek om een meer informatief studentenmodel te creëren en de modelprestaties verder te verbeteren. MMRD wordt geëvalueerd op de bekende Twitter-datasets en zou goede detectieresultaten kunnen opleveren in het vroege stadium van geruchtenverspreiding.

We introduceren participant-level rumor detection model (PLRD) en multi-view learning with attention for rumor detection model (UMLARD) in Hoofdstuk 7. Zowel PLRD als UMLARD zijn ontworpen op participant-level. PLRD maakt gebruik van verschillende fijnmazige gebruikersfuncties van de diffusiethreads, i.e., de sociale homofilie, invloed, gevoeligheid, temporele kenmerken van de gebruikers, en gebruikt deze kenmerken vervolgens om te bepalen of de informatie waar of onwaar is. PLRD introduceert ook een Variational Autoencoder (VAE) om de onzekerheid aan te pakken die bestaat in de leerfase van functies. UMLARD breidt PLRD uit en lost één brandbeperking op die is achtergelaten door PLRD, i.e., invoerfuncties verstrengeld met geleerde functies op hoog niveau, door gebruik te maken van drie view-specific inbeddingsmethoden met verschillende invoer. UMLARD stelt ook op innovatieve wijze een capsule-based attention layer voor ter vervanging van het oorspronkelijke attention mechanism in PLRD, dat effectiever is in zowel prestaties als tijdskosten. Zowel PLRD als UMARD presteren beter dan andere non-user-level modellen, wat aantoont dat het ontwikkelen van modellen op user-level inderdaad de detectieprestaties verbetert.