Table 2: Performance comparisons on the IG_US and FB_US datasets.

	Instagram				Facebook			
Measure	`Duration`	`Social`	`Personal`	`All`	`Duration`	`Social`	`Personal`	`All`
Acc.	0.34±0.02	0.59±0.01	0.69±0.03	**0.78±0.02**	0.36±0.01	0.65±0.05	0.73±0.02	**0.83±0.02**
AUC	0.36±0.02	0.61±0.01	0.74±0.01	**0.79±0.01**	0.37±0.01	0.68±0.01	0.77±0.02	**0.84±0.01**
Micro-F1	0.42±0.02	0.71±0.01	0.78±0.04	**0.85±0.01**	0.44±0.04	0.74±0.02	0.81±0.01	**0.89±0.01**
Macro-F1	0.33±0.01	0.64±0.01	0.73±0.02	**0.85±0.01**	0.35±0.02	0.68±0.02	0.77±0.03	**0.90±0.01**

Table 3: Comparisons of SNMDD with different classification techniques.

Technique	Acc.	AUC
Single-source (FB)		
J48 Decision Tree Learning	74.4%	0.750
ℓ_1-regularized ℓ_2-loss SVM	77.6%	0.781
ℓ_2-regularized ℓ_2-loss SVM	77.9%	0.783
ℓ_1-regularized logistic regression	76.3%	0.776
ℓ_2-regularized logistic regression	76.4%	0.777
DTSVM	76.4%	0.774
TSVM	**83.1%**	**0.842**
Multi-source (FB+IG)		
CF	75.5%	0.759
Tucker	85.6%	0.872
STM	**89.7%**	**0.926**

trast, users of CR and IO prefer to use social media instead of playing games alone. Moreover, people with compulsive personality are more introverted. In contrast, people with CR usually create virtual bonds to develop pathological relationships for compensation of their (missing) offline relationship.

The parasociality, effective for detecting all SNMD types, is especially useful for detecting CR cases. For example, in our user study, we find user A, 21-year-old male, frequently posting news feeds, such as "I'm so bored :(((((...Ahhhhhh!!", and his cross-dressing photos on his Facebook timeline, more than 3 times a week, which usually get less than 5 likes. At the same time, he "likes" a large number of posts from others. SNMDD classifies him as a potential CR case and his questionnaire reveals that he constantly blocks out disturbing thoughts about life and finds himself anticipating when he goes online again.

Burst intensity and length seem to be quite useful for detecting IO cases. For example, user B, 36-year-old male, is detected as IO since the behavior of clicking "likes" fits the pattern of bursts, i.e., the median of his burst intensity is high, equal to 31. His answers to the standard questionnaire reveal that he loses sleep due to late-night access on Facebook to check others' news feeds. Through interview, user B explains that he cannot stop checking for new posts and e-mails even when all his news feeds and emails are read. Some of his friends reply him: "are you a robot? no sleep needed?!?!!", indicating that user B is indulged in finding social news.

Next, we analyze the importance of different features to our classifiers. The information gain is exploited to measure the importance of each feature. In summary, the top 5 important features overall are : 1) median of the intensity of bursts, 2) online/offline interaction ratio, 3) parasociality, 4) number of used stickers, and 5) standard deviation of the length of bursts. It is worth noting that TSVM using only these 5 features in SNMDD achieves an accuracy of 76.4% and 80.7% for IG_US and FB_US, respectively, close to that of using all features (`All`). In other words, integrating important social and personal features provides good results because effective

Table 4: Top features and Acc. on the FB_US dataset.

CR	NC	IO
Parasociality	Game posts	Median of BI
Median of BI	Online/offline ratio	Online/offline ratio
Sticker number	Parasociality	SD of BL
Online/offline ratio	Number of selfies	Sticker number
CC.	CC.	Parasociality
Acc.: 80.2%	Acc.: 76.8%	Acc.: 82.7%

Table 5: Feature effectiveness analysis: SNMDD accuracy on the FB_US dataset.

Used Features	Accuracy	Used Features	Accuracy
PR	56.9%	`All`–PR	78.2%
ONOFF	60.3%	`All`–ONOFF	75.1%
SC	40.1%	`All`–SC	78.8%
SSB	44.4%	`All`–SSB	79.3%
SD	58.9%	`All`–SD	73.2%
TEMP	67.5%	`All`–TEMP	68.1%
UT	36.4%	`All`–UT	82.6%
DIS	54.0%	`All`–DIS	75.9%
PROF	18.2%	`All`–PROF	81.5%
`All`	83.1%		

personal features, e.g., the temporal behavior features, can be used to differentiate the users suffering from withdraw or relapse symptoms and heavy users, while social features capture the interactions among users to differentiate different SNMDs.

Finally, we carefully examine the effectiveness of each feature on the FB_US dataset. Table 5 compares the performance of different feature combinations using TSVM, where `All`–X means all features *excluding* category X. Unsurprisingly, combining all features leads to the best performance with the accuracy of 83.1%. In terms of individual category of features, the temporal behavior features are the most effective, whereas the profile features are the least effective. The accuracy obtained by excluding one category of features is at least 68.1%, which shows that the features are generally robust even when one feature set is missing. The accuracy of `All`–PROF is close to `All`, indicating that PROF is the least important.

Figs. 2(a) and 2(b) show the improvement of adding different features in TSVM on the FB_US dataset and the proposed *STM* on multi-source data (i.e., FB_US and IG_US). The feature selection of TSVM is based on the information gain (the top-5 features mentioned earlier), while the tensor approach automatically extracts important latent features. We observe a diminishing return property on both figures, where the improvement becomes marginal as more features are included. Fig. 2(a) shows a power fit function ($p(x) = 0.3087x^{-1.89}$) of the curve with $R^2 = 0.9512$. The exponent -1.89 denotes that the improvement by adding n-th feature is $n^{-1.89}$ times smaller than that by adding the first feature. On the other hand, the results of the tensor-based approach in Fig. 2(b) show that the accuracy increment for adding a single feature drops faster ($p(x) = 1.06x^{-2.82}$) since the proposed *STM* can extract much more important and concise features.