

*IBM SPSS Modeler 18.2 - Guida alla
distribuzione*

IBM

Nota

Prima di utilizzare queste informazioni ed il relativo prodotto, leggere le informazioni contenute in "Informazioni particolari" a pagina 33.

Informazioni sul prodotto

Questa edizione si applica alla versione 18, release 2, livello di modifica 0 di IBM SPSS Modeler e a tutte le successive release e modifiche a meno che non sia diversamente indicato nelle nuove edizioni.

Indice

Prefazione.	v	Proprietà del lavoro IBM SPSS Modeler - Esportazione Legacy TM1	17
Capitolo 1. Osservazioni generali.	1	Proprietà del lavoro IBM SPSS Modeler - Importazione TM1	17
IBM SPSS Collaboration and Deployment Services.	1	Proprietà del lavoro IBM SPSS Modeler - Esportazione TM1	17
Collaborazione.	1	Proprietà del lavoro IBM SPSS Modeler - Importazione di Analytic Server	18
Distribuzione	2	Proprietà del lavoro IBM SPSS Modeler - Esportazione di Analytic Server	18
Architettura del sistema	2	IBM SPSS Modeler Proprietà del lavoro - Notifiche	18
IBM SPSS Collaboration and Deployment Services Repository	3	Visualizzazione dei flussi IBM SPSS Modeler	19
IBM SPSS Modeler con collaborazione	4	Codici di completamento di IBM SPSS Modeler	19
IBM SPSS Deployment Manager	4	Limitazioni dei flussi di IBM SPSS Modeler.	19
IBM SPSS Collaboration and Deployment Services Deployment Portal	5	Tipi di nodi	20
IBM SPSS Deployment Manager basato su browser	5	Esempio script: Confronto tra i modelli Campione e Sfidante	20
Server di esecuzione.	5		
Utilizzo dei file IBM SPSS Modeler	6		
Capitolo 2. Definizioni server.	7	Capitolo 4. Servizio di calcolo del punteggio	23
Aggiunta di nuove definizioni server	7	Limitazioni dei flussi di IBM SPSS Modeler.	23
Parametri del server IBM SPSS Modeler	7		
Modifica delle definizioni server.	8		
Capitolo 3. Fasi lavoro di IBM SPSS Modeler.	9	Capitolo 5. Fasi di lavoro campione/sfidante.	25
Utilizzo dei flussi di IBM SPSS Modeler	9	Panoramica campione/sfidante	25
Configurazione del server di IBM SPSS Modeler	9	Metrica di valutazione dei modelli.	25
Visualizzazione delle proprietà del lavoro IBM SPSS Modeler	9	Dipendenza dall'ordine	26
IBM SPSS Modeler Proprietà del lavoro - Generale	9	Informazioni generali	26
IBM SPSS Modeler Proprietà del lavoro - File di dati	12	Sfidanti.	27
IBM SPSS Modeler Proprietà del lavoro - Vista dati	12	Selezione degli sfidanti	28
IBM SPSS Modeler Proprietà del lavoro Origini dati ODBC	13	Sfidanti non validi	28
Proprietà del lavoro IBM SPSS Modeler - Geospaziale	14	Selezione delle sorgenti dati degli sfidanti	28
IBM SPSS Modeler Proprietà del lavoro - Parametri	14	Campione	28
IBM SPSS Modeler Proprietà del lavoro - Risultati	14	Prova del campione	29
Proprietà del lavoro IBM SPSS Modeler - Importazione Cognos	16	File di dati.	31
Proprietà del lavoro IBM SPSS Modeler - Esportazione Cognos	16	Vista dati	31
Proprietà del lavoro IBM SPSS Modeler - Importazione Legacy TM1	16	origine dati ODBC	31
		Importazione Cognos	32
		Informazioni particolari.	33
		Marchi	34
		Termini e condizioni per la documentazione del prodotto	35
		Indice analitico	37

Prefazione

IBM® SPSS Modeler è l'efficace workbench di data mining aziendale di IBM Corp.. SPSS Modeler consente alle organizzazioni di migliorare le relazioni con i clienti e con il pubblico grazie a un'analisi approfondita dei dati. Le organizzazioni potranno utilizzare le informazioni ottenute tramite SPSS Modeler per mantenere i clienti di valore, cogliere opportunità di vendite incrociate, attrarre nuovi clienti, individuare frodi, diminuire i rischi e migliorare l'offerta di servizi a livello statale.

L'interfaccia visuale di SPSS Modeler favorisce l'applicazione di una competenza di business specifica da parte degli utenti, grazie alla quale sarà possibile ottenere modelli di previsione più efficaci ed una riduzione nei tempi di sviluppo delle soluzioni. SPSS Modeler offre una vasta gamma di tecniche di creazione di modelli, quali previsione, classificazione, segmentazione ed algoritmi per l'individuazione delle associazioni. IBM SPSS Modeler Solution Publisher consente quindi di distribuire a livello aziendale i modelli creati in modo che vengano utilizzati dai responsabili dei processi decisionali oppure inseriti in un database.

Informazioni su IBM Business Analytics

Il software IBM Business Analytics fornisce informazioni complete, coerenti e accurate a cui i responsabili delle decisioni possono affidarsi per ottimizzare le prestazioni di business. Un ampio portafoglio di applicazioni di business intelligence, analisi predittiva, gestione delle prestazioni e delle strategie finanziarie e analisi offre una panoramica chiara, istantanea e interattiva delle prestazioni attuali e la possibilità di prevedere i risultati futuri. Utilizzato in combinazione con potenti soluzioni di settore, prassi consolidate e servizi professionali, questo software consente alle aziende di tutte le dimensioni di ottimizzare la produttività, automatizzare le decisioni senza problemi e fornire risultati migliori.

Come parte di questo portafoglio, il software IBM SPSS Predictive Analytics consente alle aziende di prevedere gli eventi futuri e di agire tempestivamente in modo da migliorare i risultati delle attività di business. Le aziende, gli enti governativi e le università di tutto il mondo si affidano alla tecnologia IBM SPSS perché rappresenta un vantaggio concorrenziale in termini di attrazione, retention ed aumento dei clienti, riducendo al tempo stesso le frodi e limitando i rischi. Incorporando il software IBM SPSS nelle attività quotidiane, le aziende diventano imprese in grado di effettuare previsioni e di gestire ed automatizzare le decisioni, per raggiungere gli obiettivi di business e vantaggi tangibili sulla concorrenza. Per ulteriori informazioni o per contattare un rappresentante, visitare il sito <http://www.ibm.com/spss>.

Supporto tecnico

Il supporto tecnico è a disposizione dei clienti che dispongono di un contratto di manutenzione. I clienti possono contattare il supporto tecnico per richiedere assistenza per l'utilizzo di IBM Corp. o per l'installazione di uno degli ambienti hardware supportati. Per contattare il supporto tecnico, visitare il sito Web IBM Corp. all'indirizzo <http://www.ibm.com/support>. Ricordare che durante la richiesta di assistenza sarà necessario fornire i dati di identificazione personali, i dati relativi alla propria società e il numero del contratto di manutenzione.

Capitolo 1. Osservazioni generali

IBM SPSS Collaboration and Deployment Services

IBM SPSS Collaboration and Deployment Services è un'applicazione a livello aziendale che consente la diffusione dell'utilizzo e della distribuzione di analisi predittive.

IBM SPSS Collaboration and Deployment Services fornisce l'archiviazione centralizzata, protetta e verificabile degli asset analitici e funzionalità avanzate per la gestione ed il controllo di processi analitici predittivi nonché meccanismi sofisticati per la consegna dei risultati dell'elaborazione analitica agli utenti. I vantaggi di IBM SPSS Collaboration and Deployment Services comprendono:

- protezione del valore degli asset analitici
- Conformità ai requisiti legislativi
- Miglioramento della produttività degli analisti
- Riduzione al minimo dei costi IT per la gestione analitica

IBM SPSS Collaboration and Deployment Services consente di gestire in tutta sicurezza vari asset analitici e favorisce una maggiore collaborazione tra coloro che li sviluppano e li utilizzano. Inoltre, le funzionalità di distribuzione garantiscono che gli utenti acquisiscano le informazioni di cui necessitano per compiere tempestivamente le opportune azioni.

Collaborazione

Collaborazione significa capacità di condividere e riutilizzare gli asset analitici in modo efficiente e costituisce la chiave per lo sviluppo e l'implementazione delle analisi all'interno di un'azienda.

Gli analisti necessitano di una posizione in cui collocare i file che devono essere resi disponibili ad altri analisti o utenti di business. Questa posizione necessita dell'implementazione di un controllo di versione per i file, in modo che sia possibile gestire l'evoluzione dell'analisi. È necessaria protezione per controllare l'accesso e la modifica dei file. Infine, è indispensabile un meccanismo di backup e ripristino per proteggere l'azienda dalla perdita degli asset fondamentali.

Per soddisfare queste esigenze, IBM SPSS Collaboration and Deployment Services fornisce un repository per archiviare gli asset che utilizzano una gerarchia di cartelle simile alla maggior parte dei file system most file systems. I file archiviati nel IBM SPSS Collaboration and Deployment Services Repository sono disponibili per gli utenti di tutta l'azienda, a condizione che tali utenti abbiano le autorizzazioni appropriate per l'accesso. Per assistere gli utenti nella ricerca degli asset, il repository offre una funzionalità di ricerca.

Gli analisti possono lavorare con i file nel repository dalle applicazioni client che sfruttano l'interfaccia di servizio di IBM SPSS Collaboration and Deployment Services. I prodotti come IBM SPSS Statistics e IBM SPSS Modeler consentono l'interazione diretta con i file nel repository. Un analista può archiviare una versione di un file in sviluppo, recuperare questa versione in un secondo momento e continuare a modificare il file finché non è finito e pronto per essere spostato in un processo di produzione. Questi file possono includere interfacce personalizzate che eseguono processi analitici e consentono in questo modo agli utenti di business di sfruttare il lavoro dell'analista.

L'utilizzo del repository protegge l'azienda fornendo una posizione centrale per gli asset analitici consentendo di eseguire facilmente il backup ed il ripristino. Inoltre, le autorizzazioni a livello etichetta versione, file e utente controllano l'accesso ai singoli asset. Il controllo della versione e le etichette della versione oggetto garantiscono che si stanno utilizzando le versioni corrette degli asset nei processi di produzione. Infine, le funzioni di accesso offrono la possibilità di tracciare le modifiche del sistema e dei file.

Distribuzione

Per raggiungere appieno lo scopo delle analisi predittive, gli asset analitici devono fornire input per le decisioni di business. Deployment colma il divario tra l'analisi e l'azione offrendo risultati alle persone e processi secondo una pianificazione o in tempo reale.

In IBM SPSS Collaboration and Deployment Services, singoli file archiviati nel repository possono essere inclusi nell'elaborazione **lavori**. I lavori definiscono una sequenza di esecuzione per le risorse analitiche e possono essere creati con IBM SPSS Deployment Manager. I risultati dell'esecuzione possono essere archiviati nel repository, in un file system oppure inviati a destinatari specifici. Qualsiasi utente che abbia autorizzazioni sufficienti può accedere ai risultati archiviati nel repository utilizzando l'interfaccia IBM SPSS Collaboration and Deployment Services Deployment Portal. I lavori possono essere attivati in base a una pianificazione definita oppure in risposta a eventi di sistema.

Inoltre, il servizio di calcolo del punteggio di IBM SPSS Collaboration and Deployment Services consente di fornire i risultati analitici dei modelli sottoposti a deployment in tempo reale quando si interagisce con un cliente. Un modello analitico configurato per il calcolo del punteggio può combinare i dati raccolti da un'interazione corrente con il cliente con dati cronologici per produrre un punteggio che determina il corso dell'interazione. Il servizio può essere utilizzato da qualsiasi applicazione client e consente la creazione di interfacce personalizzate per la definizione del processo.

Le funzionalità di distribuzione di IBM SPSS Collaboration and Deployment Services sono pensate per integrarsi facilmente con l'infrastruttura aziendale. La funzione Single sign-on (SSO) riduce la necessità di fornire manualmente le credenziali nelle varie fasi del processo. Inoltre, il sistema può essere configurato per essere conforme allo standard Federal Information Processing Standard Publication 140-2.

Nota: Se un flusso SPSS Modeler contiene un nodo che utilizza un tipo di elenco, il ramo contenente quel nodo non supporta il servizio di calcolo del punteggio.

Architettura del sistema

In generale, IBM SPSS Collaboration and Deployment Services è costituito da un singolo IBM SPSS Collaboration and Deployment Services Repository centralizzato che serve vari client e utilizza server di esecuzione per elaborare gli asset analitici.

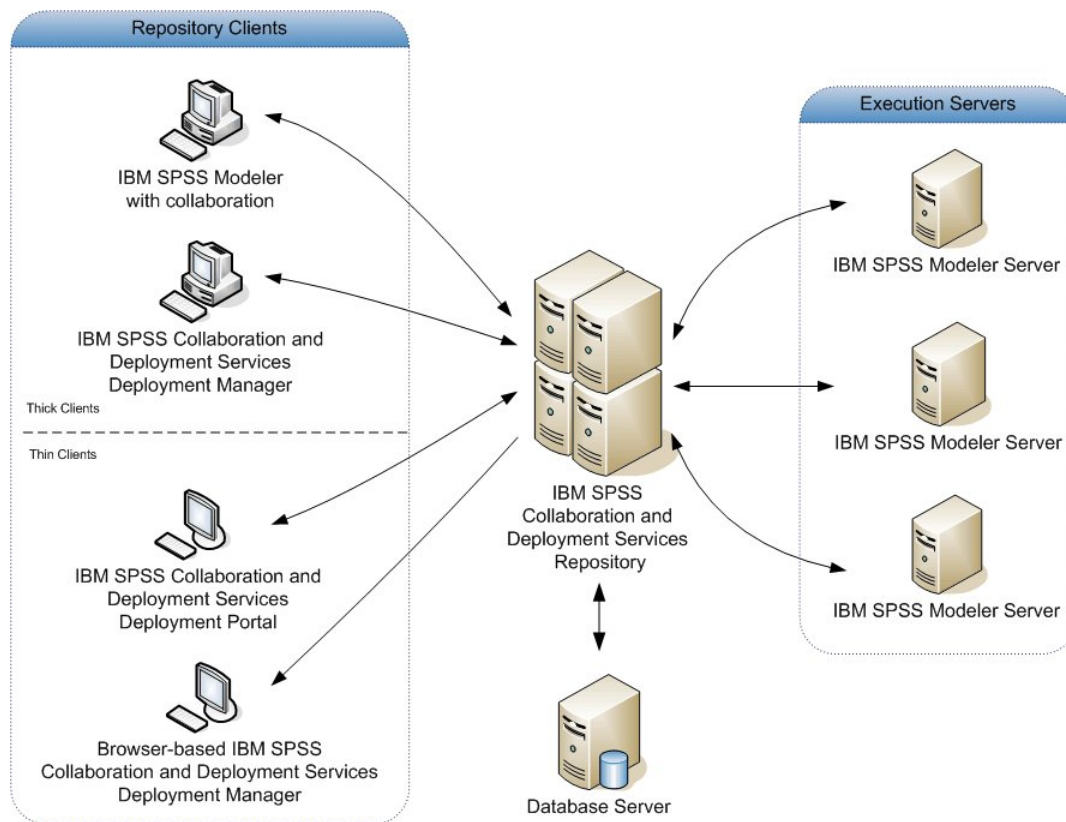


Figura 1. Architettura IBM SPSS Collaboration and Deployment Services

IBM SPSS Collaboration and Deployment Services è costituito dai componenti seguenti:

- IBM SPSS Collaboration and Deployment Services Repository per i dati analitici
- IBM SPSS Modeler
- IBM SPSS Deployment Manager
- IBM SPSS Collaboration and Deployment Services Deployment Portal
- IBM SPSS Deployment Manager basato su browser

IBM SPSS Collaboration and Deployment Services Repository

Il repository fornisce una posizione centralizzata per archiviare gli asset analitici come ad esempio modelli e dati. Il repository richiede l'installazione di un database relazionale come ad esempio IBM Db2, Microsoft SQL Server o Oracle.

Il repository comprende le funzionalità per:

- Sicurezza
- Controllo della versione
- Ricerca
- Esplorazione

Le opzioni di configurazione per il repository vengono definite utilizzando IBM SPSS Deployment Manager o IBM SPSS Deployment Manager basato sul browser. I contenuti del repository vengono gestiti con Deployment Manager e che è possibile accedere mediante IBM SPSS Collaboration and Deployment Services Deployment Portal.

IBM SPSS Modeler con collaborazione

IBM SPSS Modeler con collaborazione consente l'interazione con il IBM SPSS Collaboration and Deployment Services Repository dall'interno IBM SPSS Modeler . I file possono essere archiviati e recuperati direttamente da IBM SPSS Modeler .

Inoltre i file di sintassi IBM SPSS Modeler i flussi archiviati nel repository possono essere eseguiti come fasi all'interno dei lavori. Un lavoro può contenere un numero qualsiasi di fasi, ciascuna corrispondente a un file separato. Le relazioni definite tra le fasi determinano il flusso di elaborazione. Il lavoro può essere pianificato per essere eseguito a un orario specifico, in base a uno schema di ricorrenza o in risposta a un evento definito. Inoltre, possono essere inviate notifiche a destinatari specifici per informarli dello stato di esecuzione di una singola fase e dell'intero lavoro.

La collaborazione tra IBM SPSS Collaboration and Deployment Services e IBM SPSS Modeler è consentita mediante l'utilizzo di adattatori. Questi adattatori vengono installati nell'ambiente IBM SPSS Collaboration and Deployment Services per aggiungere le funzioni specifiche del prodotto. Per ulteriori informazioni, consultare la documentazione IBM SPSS Modeler .

Nota: Se si dispone di un flusso SPSS Modeler che utilizza il nodo sorgente di Analytic Server, è necessario consentire una connessione diretta tra il IBM SPSS Collaboration and Deployment Services Server e l'Analytic Server. Altrimenti, il job fallirà se il firewall bloccherà la connessione tra i due server.

IBM SPSS Deployment Manager

IBM SPSS Deployment Manager è un'applicazione client per IBM SPSS Collaboration and Deployment Services Repository che consente agli utenti di pianificare, automatizzare ed eseguire attività analitiche come l'aggiornamento di modelli o la generazione di punteggi.

L'applicazione client consente ad un utente di eseguire le seguenti attività:

- Vedere qualsiasi file esistente all'interno del sistema, compresi report IBM SPSS Modeler stream, scenari, e file di dati
- Importare i file nel repository
- Pianificare i lavori da eseguire ripetutamente utilizzando un pattern di ricorrenza specificato come ogni trimestre o ogni ora.
- Modificare le proprietà del lavoro esistenti
- Determinare lo stato di un lavoro
- Specificare la notifica email dello stato del lavoro

Inoltre, l'applicazione client consente agli utenti di eseguire attività amministrative per IBM SPSS Collaboration and Deployment Services, compreso:

- Gestire utenti
- Configurare i provider di sicurezza
- Assegnare ruoli ed azioni

IBM SPSS Deployment Manager basato su browser

IBM SPSS Deployment Manager basato su browser è un'interfaccia thin-client che consente di effettuare le attività di configurazione e gestione del sistema seguenti:

- Impostazione delle opzioni di configurazione del sistema
- Configurazione dei provider di sicurezza
- Gestione dei tipi MIME

Gli utenti non amministrativi possono eseguire tutte queste attività, a condizione che le azioni associate alle credenziali di accesso siano appropriate. Le azioni vengono assegnate da un amministratore.

Generalmente accedere a IBM SPSS Deployment Manager basato su browser al seguente URL:
`http://<host IP address>:<port>/security/login`

Nota: Un indirizzo IPv6 deve essere compreso tra parentesi quadre, come ad esempio [3ffe:2a00:100:7031::1].

Se l'ambiente è configurato per utilizzare un percorso di contesto personalizzato per le connessioni al server, includere tale percorso nell'URL.

`http://<host IP address>:<port>/<context path>/security/login`

IBM SPSS Collaboration and Deployment Services Deployment Portal

IBM SPSS Collaboration and Deployment Services Deployment Portal è un'interfaccia thin-client per accedere al repository. Diversamente da IBM SPSS Deployment Manager basato su browser che è destinato agli amministratori IBM SPSS Collaboration and Deployment Services Deployment Portal è un portale Web utilizzato da diversi utenti.

Il portale Web include le seguenti funzionalità:

- Ricerca del contenuto del repository per cartella
- Apertura del contenuto pubblicato
- Esecuzione di lavori
- Generazione di punteggi utilizzando modelli archiviati nel repository
- Ricerca del contenuto del repository
- Visualizzazione delle proprietà del contenuto
- Accesso alle singole preferenze utente, come l'indirizzo email e la password, opzioni generali, abbonamenti e opzioni per i formati del file di output.

Generalmente si accede alla home page al seguente URL:

`http://<host IP address>:<port>/peb`

Nota: Un indirizzo IPv6 deve essere compreso tra parentesi quadre, come ad esempio [3ffe:2a00:100:7031::1].

Se l'ambiente è configurato per utilizzare un percorso di contesto personalizzato per le connessioni al server, includere tale percorso nell'URL.

`http://<host IP address>:<port>/<context path>/peb`

IBM SPSS Deployment Manager basato su browser

IBM SPSS Deployment Manager basato su browser è un'interfaccia thin-client che consente di effettuare le attività di configurazione e gestione del sistema seguenti:

- Configurazione del sistema
- Configurazione dei provider di protezione
- Gestione dei tipi MIME

Gli utenti non amministrativi possono eseguire tutte queste attività, a condizione che le azioni associate alle credenziali di accesso siano appropriate. Le azioni vengono assegnate da un amministratore.

Server di esecuzione

I server di esecuzione forniscono la possibilità di eseguire le risorse archiviate all'interno del repository. Quando una risorsa viene inclusa in un lavoro per l'esecuzione, per definire la fase del lavoro occorre anche specificare il server di esecuzione utilizzato per l'elaborazione della fase. Il tipo di server di esecuzione dipende dalla risorsa.

Il server di esecuzione IBM SPSS Modeler è IBM SPSS Modeler Server. Esso consente l'analisi distribuita per il data mining e la creazione del modello. Questo server di esecuzione richiede di specificare le credenziali utente con le quali viene effettuata l'elaborazione. Consultare Capitolo 2, "Definizioni server", a pagina 7 per ulteriori informazioni.

Per consentire il bilanciamento del carico, due o più server di esecuzione possono essere raggruppati in un cluster di server. Quando una fase di un lavoro utilizza un cluster per l'esecuzione, IBM SPSS Collaboration and Deployment Services determina quale server gestito all'interno del cluster sia più adatto per gestire i requisiti di elaborazione in quel momento. Per ulteriori informazioni, consultare la documentazione di IBM SPSS Deployment Manager.

Durante la creazione del lavoro, assegnare un server di esecuzione o un cluster di server a ciascuna fase inclusa nel lavoro. Quando si esegue il lavoro, il repository utilizza i server di esecuzione specificati per eseguire le corrispondenti analisi.

Utilizzo dei file IBM SPSS Modeler

In IBM SPSS Deployment Manager, in generale per utilizzare i file di IBM SPSS Modeler occorre:

1. Definire un server di esecuzione per elaborare IBM SPSS Modeler il flusso o lo scenario se non esiste già un'appropriata definizione.
2. Aggiungere il IBM SPSS Modeler flusso o scenario ad un lavoro come fase lavoro.
3. Specificare le proprietà del lavoro incluso le pianificazioni per l'esecuzione.

Capitolo 2. Definizioni server

L'esecuzione di IBM SPSS Collaboration and Deployment Services Repository resource come fase di processo richiede la specifica di un appropriato server corrispondente per elaborare le istruzioni contenute nella fase del processo. Le informazioni relative alla connessione per tale server vengono specificate all'interno della definizione del server.

Le definizioni server possono essere classificate come server di esecuzione o server repository.

- I server di esecuzione elaborano il contenuto di una risorsa IBM SPSS Collaboration and Deployment Services Repository. Il tipo di server di esecuzione deve corrispondere al tipo di risorsa elaborato. Un IBM SPSS Modeler fase di lavoro richiede un IBM SPSS Modeler definizione del server.
- Un server repository corrisponde a un'installazione repository IBM SPSS Collaboration and Deployment Services. Un server di questo tipo viene utilizzato generalmente dalle fasi dei lavori che necessitano di restituire i risultati a un repository.

Le definizioni server sono contenute nella cartella *Definizioni risorse* di Esplora contenuto. Nello specifico, sono definite nella sottocartella *Server*.

Aggiunta di nuove definizioni server

Per aggiungere un nuovo server:

1. In Esplora contenuto, aprire la cartella *Definizioni risorse*.
2. fare clic sulla cartella *Server*.
3. Dal menu File scegliere:

Nuovo > Definizione server

Viene visualizzata la procedura guidata di aggiunta di una nuova definizione server. In alternativa, è possibile accedere alla finestra di dialogo della nuova definizione del server facendo clic su **Nuovo** accanto al campo server della scheda generale per alcune fasi. La procedura per la definizione di nuovi server è la seguente:

1. Assegnare un nome alla definizione server e specificarne il tipo. Si noti che i tipi di server disponibili dipendono da quali adattatori di prodotto sono installati nel repository.
2. Selezionare una posizione per la definizione nella cartella *Server*.
3. Specificare i parametri per il server che definiscono le informazioni di esecuzione o connessione. Il set di parametri dipende dal tipo di server.

Parametri del server IBM SPSS Modeler

Una definizione server IBM SPSS Modeler specifica i parametri di connessione per i server IBM SPSS Modeler utilizzati per elaborare le fasi dei lavori. I flussi di IBM SPSS Modeler vengono eseguiti sul server IBM SPSS Modeler.

1. Nel campo Host, immettere il nome dell'host in cui risiede il server. Ad esempio, se si crea una definizione server IBM SPSS Modeler, l'host è il computer che contiene il server IBM SPSS Modeler.
2. Nel campo Porta, immettere il numero di porta da utilizzare per la connessione all'host.
3. Nel campo Percorso dati di default, immettere il percorso in cui collocare i file di dati.
4. Se è necessario utilizzare il protocollo SSL (Secure Socket Layer) per la connessione server, selezionare **Questa è una porta protetta**.
5. Fare clic su **Fine**. La nuova definizione viene visualizzata nella cartella *Server*.

Nota: Quando si esegue un lavoro IBM SPSS Collaboration and Deployment Services in un flusso Valutazione, viene creato un file temporaneo. Per impostazione predefinita, il file viene salvato nella directory di installazione di IBM SPSS Modeler Server. È possibile modificare la cartella di dati predefinita in cui vengono salvati i dati durante la creazione della connessione a IBM SPSS Modeler Server in IBM SPSS Modeler.

Modifica delle definizioni server

Per modificare una definizione server:

1. In Esplora contenuto, aprire la cartella *Definizioni risorse*.
2. Aprire la cartella *Server*.
3. Fare doppio clic sul server da modificare. Viene visualizzata la finestra di dialogo Modifica definizione server.
4. Modificare i parametri della definizione server come necessario.
5. Fare clic su Fine per salvare le modifiche.

Capitolo 3. Fasi lavoro di IBM SPSS Modeler

Utilizzo dei flussi di IBM SPSS Modeler

Questa sezione descrive i flussi di IBM SPSS Modeler nel contesto di IBM SPSS Deployment Manager. IBM SPSS Modeler i flussi sono presenti in Deployment Manager completamente formato. Per informazioni dettagliate su come creare e utilizzare i flussi, consultare la documentazione di IBM SPSS Modeler.

È possibile utilizzare i flussi di IBM SPSS Modeler in Deployment Manager. Come in ogni altra fase, un flusso di IBM SPSS Modeler deve essere aggiunto a un lavoro prima di poter essere eseguito mediante Deployment Manager. Nello specifico, è possibile effettuare le attività riportate di seguito:

- Importare i flussi
- Modificare i parametri dei flussi
- Eseguire i flussi

Nota: Se si dispone di un flusso SPSS Modeler che utilizza il nodo sorgente di Analytic Server, è necessario consentire una connessione diretta tra il IBM SPSS Collaboration and Deployment Services Server e l'Analytic Server. Altrimenti, il job fallirà se il firewall bloccherà la connessione tra i due server.

Configurazione del server di IBM SPSS Modeler

Prima di iniziare a utilizzare i flussi di IBM SPSS Modeler in Deployment Manager, è necessario effettuare le seguenti attività di configurazione:

- Creare una definizione server di IBM SPSS Modeler. Consultare l'argomento "Parametri del server IBM SPSS Modeler" a pagina 7 per ulteriori informazioni.
- Definire le credenziali del server.

Visualizzazione delle proprietà del lavoro IBM SPSS Modeler

Quando si seleziona un flusso IBM SPSS Modeler all'interno di un lavoro, vengono visualizzate le seguenti proprietà del lavoro:

- Generale
- File di dati
- origine dati ODBC
- Parametri
- Risultati
- Importazione Cognos
- Esportazione Cognos
- Notifiche

IBM SPSS Modeler Proprietà del lavoro - Generale

Per default, la scheda delle proprietà generali viene visualizzata quando si fa clic su un flusso di IBM SPSS Modeler nell'area di lavoro.

La scheda delle proprietà generali contiene le informazioni seguenti:

Nome fase lavoro. Il nome della fase di lavoro. Generalmente, il nome della fase del lavoro è costituito dal nome del flusso di IBM SPSS Modeler e dal suffisso `_step`. Tuttavia, è possibile modificare questo nome.

Per modificare il nome della fase del lavoro, digitare il nuovo nome nel campo **Nome della fase del lavoro**. Le modifiche vengono riportate nell'area del lavoro.

IBM SPSS Modeler Flusso. Il nome e il percorso del flusso originale di IBM SPSS Modeler. Lo stream sorgente non è modificabile.

Sovrascrivi tipo Utilizzare questa opzione per specificare se le impostazioni in una fase di lavoro o un nodo in un flusso hanno priorità durante l'esecuzione.

- **Il lavoro sovrascrive il flusso** Se viene selezionata questa opzione, quando viene eseguito il lavoro, le proprietà del nodo utilizzano i valori impostati nella fase del lavoro.
- **Il flusso sovrascrive il lavoro** Selezionando questa opzione, i controlli della proprietà del nodo vengono visualizzati nella fase del lavoro; le uniche eccezioni sono alcuni controlli delle credenziali. Se vengono modificate le proprietà del nodo nel flusso, quando viene eseguita la fase del lavoro, vengono utilizzate le proprietà del nodo modificate.

Le opzioni **Sovrascrivi tipo** hanno effetto sulle impostazioni delle credenziali nei seguenti modi:

- Se si seleziona **Il lavoro sovrascrive il flusso**, è possibile modificare i controlli di selezione delle credenziali e le credenziali selezionate diventano effettive al runtime.
- Se si seleziona **Il flusso sovrascrive il lavoro**, sono possibili due casi:
 - Quando vengono creati i nodi nel client SPSS Modeler che utilizza le credenziali, la modalità di autenticazione è **Credenziali archiviate** ed il nome delle credenziali esiste in IBM SPSS Collaboration and Deployment Services, il controllo di selezione delle credenziali non è disponibile. Tuttavia, se il nome delle credenziali non esiste in IBM SPSS Collaboration and Deployment Services, è possibile utilizzare il controllo di selezione delle credenziali; le credenziali selezionate diventano effettive al runtime.
 - Quando vengono creati i nodi nel client SPSS Modeler che utilizza le credenziali, se la modalità di autenticazione è **Nome utente e password**, è possibile utilizzare il controllo di selezione delle credenziali; le credenziali selezionate diventano effettive al runtime.

Versione oggetto. La versione etichettata del flusso di IBM SPSS Modeler da utilizzare. Dall'elenco a discesa **Versione oggetto**, selezionare la versione etichettata da utilizzare.

Elenco variabili iterative. Se la fase funge da consumatore iterativo, identificare la variabile del produttore iterativo che fornisce i valori per la fase. La fase del produttore iterativo deve essere visualizzata immediatamente prima della fase nel lavoro affinché questa fase possa fungere da consumatore iterativo.

Server IBM SPSS Modeler. Server o cluster di server IBM SPSS Modeler su cui verrà eseguito il flusso. L'elenco contiene tutti i server e i cluster di server attualmente configurati per eseguire fasi di IBM SPSS Modeler. Per modificare il server, selezionarlo dall'elenco a discesa **IBM SPSS Modeler Server**. Per creare una nuova definizione server, fare clic su **Nuovo** per avviare la procedura guidata di definizione server.

IBM SPSS Modeler accesso. Credenziali utilizzate per accedere al server o al cluster di server di IBM SPSS Modeler. Per modificare le credenziali, selezionare una definizione di credenziale dall'elenco a discesa **IBM SPSS Modeler Login**. Per definire nuove credenziali, fare clic su **Nuovo** per avviare la procedura guidata di definizione delle credenziali.

Content Repository Server. Il Content Repository Server consente a un lavoro di salvare i file in un IBM SPSS Collaboration and Deployment Services Repository. Generalmente, il Content Repository Server viene specificato quando si aggiornano i modelli mediante IBM SPSS Modeler. Per specificare un Content Repository Server, selezionare un server dall'elenco a discesa **Content Repository Server**. Per creare una nuova definizione server, fare clic su **Nuovo** per avviare la procedura guidata di definizione server. Per

generare una definizione Content Repository Server in base alle informazioni server correnti, fare clic su **Genera**. Viene creata e popolata automaticamente una definizione server nel campo *Content Repository Server*.

Accesso a repository di contenuti Informazioni di accesso per il Content Repository Server. Per specificare le informazioni di accesso al repository di contenuti, selezionare una credenziale dall'elenco a discesa **Accesso a repository di contenuti**. Per creare un nuovo accesso, fare clic su **Nuovo** per avviare la procedura guidata di accesso al repository di contenuti. Se non viene utilizzata la funzione Single sign-on (SSO) per la connessione al IBM SPSS Collaboration and Deployment Services Repository, fare clic su **Genera** per generare un accesso per il Content Repository Server basato sulle impostazioni di protezione esistenti. Viene creato e popolato automaticamente un accesso per il repository di contenuti nel campo *Accesso a repository di contenuti*. La generazione dell'accesso non è disponibile quando si utilizza la funzione Single sign-on (SSO).

Espressione di avviso. Definisce gli avvisi per le fasi dei lavori collegate da un connettore Condizionale. L'espressione di avviso (per esempio, `completion_code`, `warning` o `success`) deve essere in caratteri minuscoli.

Per utilizzare le espressioni di avviso:

1. Collegare due fasi di un lavoro con un connettore Condizionale. Nel campo **Espressione** del connettore condizionale, digitare `warning==true`.
2. Passare alla scheda Generale della fase lavoro padre.
3. Nel campo **Espressione di avviso** specificare un codice di avviso, per esempio, `completion_code==18`. Questa espressione ha la precedenza sull'eventuale codice di avviso predefinito.

Quando viene eseguito il lavoro, il sistema esegue la fase lavoro padre, quindi valuta la condizione di `warning==true`. Se la condizione è `true` (vero), il sistema esamina l'espressione di avviso specificata e determina se la condizione è stata soddisfatta. Se la condizione specificata nell'espressione di avviso è stata soddisfatta, il sistema passa alla fase successiva del lavoro.

Tipo. Il modo in cui si desidera eseguire il flusso.

- **Esegui flusso.** Con questa opzione è possibile eseguire il flusso completo oppure, se il flusso contiene rami, scegliere uno o più rami da eseguire. Nel secondo caso, viene visualizzato l'elenco delle **opzioni di esecuzione**. Per default vengono selezionati tutti i rami, quindi tutti i rami vengono eseguiti quando si esegue il flusso. Per migliorare le prestazioni, deselezionare eventuali rami che non si desidera eseguire.

Nota: Se il flusso contiene uno script ed è configurato per "Esegui script" al momento della sua esecuzione, nessun ramo di esecuzione verrà visualizzato nell'elenco nell'editor di lavori di IBM SPSS Collaboration and Deployment Services.

- **Gestione dei modelli.** Scegliere questa opzione se si desidera che il flusso utilizzi le funzioni di gestione dei modelli, come la valutazione, l'aggiornamento e il calcolo del punteggio. Scegliere la funzione dall'elenco dei tipi visualizzato. **Opzioni di valutazione**

Prestazioni. Specificare le percentuali di soglia per le seguenti categorie: Cattiva o Buona. La percentuale della categoria Migliore non viene definita dall'utente. Il sistema sottrae la percentuale Buona da 100 per determinare la soglia Migliore. I valori di soglia sono obbligatori per i flussi di valutazione.

Metriche. I criteri di misurazione in base ai quali viene valutata l'efficacia del flusso. I valori validi sono Accreditamento, Precisione e Guadagni. Se si seleziona **Guadagni**, è necessario specificare anche un percentile. Se si seleziona **Accreditamento**, è necessario specificare la fase di accreditamento. I valori validi includono "Raccogli statistiche" ed "Esegui valutazione".

Opzioni di aggiornamento

Mostra un elenco di tutte le relazioni del nugget del modello nel flusso che possono essere aggiornate utilizzando i nomi dei nodi definiti nel flusso. Scegliere una relazione specifica da aggiornare

selezionando la casella visualizzata prima della relazione. Per impedire che una relazione venga aggiornata durante l'esecuzione della fase, è necessario deselegionarla. Per selezionare tutte le relazioni per l'aggiornamento, fare clic sul pulsante Seleziona tutto. Per deselegionare tutte le relazioni, fare clic sul pulsante Deseleziona tutto.

Opzioni di calcolo del punteggio

Scegliere il ramo o i rami da designare come rami di calcolo del punteggio.

Linee guida per la fornitura IBM SPSS Modeler e informazioni sul repository dei contenuti

Per tutte le fasi lavoro di IBM SPSS Modeler, sono richieste le informazioni del server IBM SPSS Modeler e le informazioni di accesso di IBM SPSS Modeler.

IBM SPSS Modeler Proprietà del lavoro - File di dati

La tabella File di dati contiene le informazioni seguenti:

Nome nodo. Nome del nodo origine che contiene i dati utilizzati dal flusso. Il nome è preceduto da un prefisso composto dal nome di ogni supernodo contenente il nodo, racchiuso tra barre. Ad esempio, se il nodo *MioNodo* è contenuto in un supernodo denominato *Supernodo1*, il nome viene visualizzato in questo modo */Supernodo1/MioNodo*. Il nome del nodo non è modificabile.

Tipo di nodo. Tipo di nodo definito nel flusso. Il tipo di nodo non è modificabile.

Nome file. Nome del file di dati di input. Per modificarlo, fare clic nella cella Nome file e modificare il nome.

Formato. Formato del file di output, per esempio un file delimitato da virgole. Per modificare il tipo di formato del file, fare clic nella cella Formato. Viene visualizzato un elenco a discesa. Selezionare il tipo di formato.

Posizione. Posizione dei file di dati di input. Per modificare la posizione, fare clic nella colonna e quindi sul pulsante con i puntini di sospensione così visualizzato. Viene visualizzata la finestra di dialogo Posizione file di input. Modificare la posizione in funzione delle proprie esigenze.

I nodi all'interno di supernodi bloccati non sono accessibili. Non possono essere né visualizzati né modificati.

IBM SPSS Modeler Proprietà del lavoro - Vista dati

Se il flusso contiene un nodo Geospaziale, in questo punto vengono visualizzati i dettagli della connessione.

Nome nodo. Il nome del nodo Vista dati.

Vista dati analitici. La vista dati analitici utilizzata.

Etichetta. L'etichetta utilizzata.

Nome tabella. Il nome della tabella di database utilizzata.

Piano di accesso ai dati. Selezionare un piano di accesso ai dati dalla vista di dati analitici. Un piano di accesso ai dati associa le tabelle dei modelli di dati in una vista di dati analitici alle origini dati fisiche. Una vista di dati analitici in genere contiene più piani di accesso ai dati. Quando si modifica il piano di accesso ai dati in uso, si modificano i dati utilizzati dal flusso. Ad esempio, se la vista di dati analitici

contiene un piano di accesso ai dati per l'addestramento di un modello e un piano di accesso ai dati per la verifica di un modello, è possibile passare dai dati di addestramento a quelli di verifica cambiando il piano di accesso ai dati in uso.

IBM SPSS Modeler Proprietà del lavoro Origini dati ODBC

La tabella *Origini dati ODBC* contiene le seguenti informazioni

Nome nodo. Nome del nodo origine che contiene i dati utilizzati dal flusso. Il nome è preceduto da un prefisso composto dal nome di ogni supernodo contenente il nodo, racchiuso tra barre. Ad esempio, se il nodo *MioNodo* è contenuto in un supernodo denominato *Supernodo1*, il nome viene visualizzato in questo modo */Supernodo1/MioNodo*.

Tipo di nodo. Tipo di nodo definito nel flusso.

Origini dati ODBC. Il nome corrente della sorgente dati ODBC (DSN). Per passare ad una sorgente dati ODBC diversa, fare clic sulla cella contenente il nome della sorgente dati corrente, quindi fare clic sul "..." pulsante che viene visualizzato. In questo modo viene visualizzata una finestra di dialogo da dove è possibile scegliere un DSN esistente o crearne uno nuovo. Notare che le impostazioni della fase lavoro sovrascrivono sempre il flusso; per se si modifica il flusso per utilizzare una diversa origini dati, è necessario anche modificare il lavoro per utilizzare la stessa origine, altrimenti l'esecuzione del lavoro può avere esito negativo.

Credenziali. Per modificare il nome utente e la password del database quando si cambia la sorgente dati ODBC, fare clic sulla cella contenente le credenziali correnti, quindi fare clic sul "..." pulsante che viene visualizzato. In questo modo viene visualizzata una finestra di dialogo da dove è possibile scegliere una definizione di credenziali esistente o crearne una nuova.

Tabella di database. Tabella di database che corrisponde al nodo.

I nodi all'interno di supernodi bloccati non sono accessibili. Non possono essere né visualizzati né modificati.

Modifica della connessione ODBC

È possibile cambiare la connessione ODBC con una già definita, oppure con quella che si crea qui.

Utilizza il DSN ODBC esistente. Fare clic su **Visualizza** per visualizzare un elenco dei DSN ODBC esistenti tra cui scegliere.

Crea un nuovo DSN ODBC. Selezionare questa opzione e fare clic su **Nuovo** per definire una nuova connessione ODBC da utilizzare con le fasi di lavoro del flusso.

Ricerca di una connessione ODBC

Questo è un elenco di tutte le connessioni ODBC che sono state definite su questo host e a cui si ha accesso.

Sceglierne una dall'elenco e fare clic su **OK**.

Modifica delle credenziali del database

Quando la connessione viene modificata, è possibile cambiare in modo da utilizzare in questo punto il nome utente e la password del database corrispondente. È possibile utilizzare una definizione di credenziale esistente dal repository, oppure creare una nuova definizione.

Genera le credenziali del repository dal nome utente password esistenti. (visualizzata solo se si importa un lavoro dalla release 4.2.1 di IBM SPSS Collaboration and Deployment Services) Fare clic sul pulsante

Genera per generare le credenziali del repository da un nome utente e password esistenti, se impostati nel lavoro importato. Ad esempio, se il nome utente sa già esiste, facendo ripetutamente clic su **Genera** verranno creati i nomi utente sa1, sa2 e così via.

Utilizza credenziale esistente. Fare clic sul pulsante **Visualizza** per visualizzare un elenco delle definizioni di credenziale esistenti tra cui scegliere.

Crea nuova credenziale. Selezionare questa opzione e fare clic su **Nuovo** per creare una nuova definizione di credenziale da utilizzare con la connessione corrente.

Ricerca di una definizione di credenziale

Questo è un elenco di tutte le definizioni di credenziali che sono state create su questo host e a cui si ha accesso.

Sceglierne una dall'elenco e fare clic su **OK**.

Proprietà del lavoro IBM SPSS Modeler - Geospaziale

Se il flusso contiene un nodo Geospaziale, in questo punto vengono visualizzati i dettagli della connessione.

Tipo di origine. Il tipo di origine dati.

Nome file. Se si utilizza un file di forma, specificare il nome file.

URL del servizio mappa. Se si utilizza un servizio Mappa, specificare l'URL del servizio.

ID mappa. Specificare l'ID mappa.

IBM SPSS Modeler Proprietà del lavoro - Parametri

È possibile modificare i parametri dei flussi di IBM SPSS Modeler utilizzando IBM SPSS Deployment Manager. La tabella Parametri contiene le informazioni seguenti:

Nome. Nome del parametro.

Archiviazione. Descrive le modalità di archiviazione del parametro, per esempio come stringa.

Valore. Valore del parametro.

Tipo. Tipo di parametro.

I parametri all'interno di supernodi bloccati non sono accessibili. Non possono essere né visualizzati né modificati.

IBM SPSS Modeler Proprietà del lavoro - Risultati

Per visualizzare le impostazioni di output per una fase di un lavoro di IBM SPSS Modeler, fare clic sulla scheda Risultati.

Nome nodo. Il nome del nodo che contiene l'output dell'elaborazione del flusso. Il nome è preceduto da un prefisso composto dal nome di ogni supernodo contenente il nodo, racchiuso tra barre. Ad esempio, se il nodo *MioNodo* è contenuto in un supernodo denominato *Supernodo1*, il nome viene visualizzato in questo modo: */Supernodo1/MioNodo*. Il nome del nodo non è modificabile.

Tipo di nodo. Il tipo di nodo come è definito nel flusso. Il tipo di nodo non è modificabile.

Nome file. Il nome del file corrispondente. Per modificare il nome, fare clic nella colonna Nome file e digitare il nuovo nome.

Nota: L'estensione file viene determinata dal formato del file selezionato che viene aggiunto automaticamente al nome del file.

Formato. Formato del file di output. Per modificare il tipo di formato del file, fare clic nella colonna Formato. Viene visualizzato un elenco a discesa. Selezionare il tipo di formato.

Posizione. Posizione del file. Per modificare la posizione, aprire la finestra di dialogo Posizione risultati facendo clic nella colonna e quindi sul pulsante con i puntini di sospensione così visualizzato.

Autorizzazioni. Autorizzazioni di accesso al file se questo è salvato nel repository. Per modificare le autorizzazioni, aprire la finestra di dialogo Autorizzazioni output facendo clic nella colonna Autorizzazioni e quindi sul pulsante con i puntini di sospensione così visualizzato.

Proprietà. Proprietà (metadati) del file. Per definire le proprietà, aprire la finestra di dialogo Proprietà output facendo clic nella colonna Proprietà e quindi sul pulsante con i puntini di sospensione così visualizzato.

I nodi all'interno di supernodi bloccati non sono accessibili. Non possono essere né visualizzati né modificati.

Visualizzazione dei risultati di output.

Per visualizzare i risultati di output del flusso, fare doppio clic sui risultati che si desidera visualizzare nella colonna Risultati della tabella Cronologia fase lavoro. I risultati vengono visualizzati in una finestra separata.

Ad esempio, si supponga che uno dei risultati di output sia un istogramma. Un istogramma di esempio potrebbe avere l'aspetto seguente.

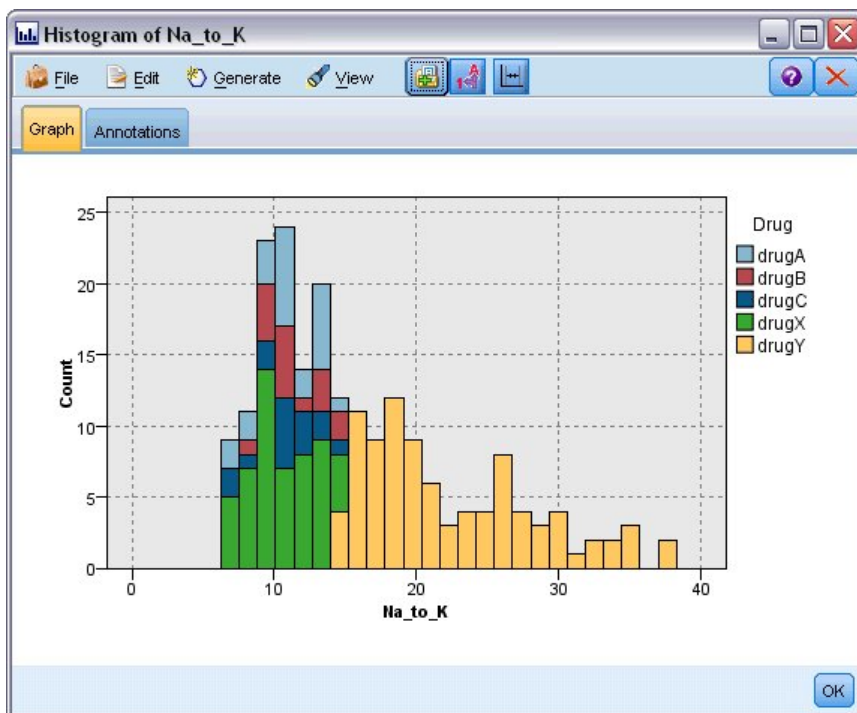


Figura 2. Istogramma di esempio

Proprietà del lavoro IBM SPSS Modeler - Importazione Cognos

Se il flusso contiene un nodo di origine IBM Cognos, vengono visualizzati i dettagli della connessione Cognos.

Nome nodo. Nome del nodo origine Cognos.

URL di connessione. URL del server Cognos al quale viene stabilita la connessione.

Nome package. Nome del pacchetto Cognos da cui vengono importati i metadati.

Anonimo. Include **Anonymous** se viene effettuato un accesso anonimo per la connessione al server Cognos oppure **Credential** se vengono utilizzati un nome utente e una password Cognos.

Credenziali. Il nome utente e la password (se richiesta) sul server Cognos.

Nota: Le credenziali Cognos devono essere create in un dominio che rappresenta l'ID dello spazio dei nomi Cognos.

Proprietà del lavoro IBM SPSS Modeler - Esportazione Cognos

Se il flusso contiene un nodo di esportazione IBM Cognos, vengono visualizzati i dettagli della connessione Cognos e ODBC.

Nome nodo. Nome del nodo di esportazione Cognos.

URL di connessione. URL del server Cognos al quale viene stabilita la connessione.

Nome package. Nome del pacchetto Cognos utilizzato per l'esportazione dei metadati.

Origine dati Nome del database Cognos utilizzato per l'esportazione dei dati.

Cartella. Percorso e nome della cartella del server Cognos dove viene creato il pacchetto di esportazione.

Anonimo. Include **Anonymous** se viene effettuato un accesso anonimo per la connessione al server Cognos oppure **Credential** se vengono utilizzati un nome utente e una password Cognos.

Credenziali. Il nome utente e la password (se richiesta) sul server Cognos.

DSN. DNS (Database Source Name) del database Cognos.

Nome tabella. Nome della tabella di database utilizzata per l'esportazione.

Credenziali. Il nome utente e la password utilizzati per la connessione al database utilizzato per l'esportazione.

Nota: Le credenziali Cognos devono essere create in un dominio che rappresenta l'ID dello spazio dei nomi Cognos.

Proprietà del lavoro IBM SPSS Modeler - Importazione Legacy TM1

Se il flusso contiene un nodo di importazione Legacy TM1, i dettagli della connessione vengono visualizzati in questo punto.

Nome nodo. Il nome del nodo di importazione Legacy TM1.

Server TM1. Il nome del server Cognos TM1.

Cubo. Il cubo TM1 da cui verranno importati i dati.

Vista. La vista da importare dal cubo TM1.

Credenziale. Credenziale del server Cognos TM1.

Proprietà del lavoro IBM SPSS Modeler - Esportazione Legacy TM1

Se il flusso contiene un nodo di esportazione Legacy TM1, i dettagli della connessione vengono visualizzati in questo punto.

Nome nodo. Il nome del nodo di esportazione Legacy TM1.

Server TM1. Il nome del server Cognos TM1.

Cubo. Il cubo TM1 in cui verranno esportati i dati.

Credenziale. Credenziale del server Cognos TM1.

Proprietà del lavoro IBM SPSS Modeler - Importazione TM1

Se il flusso contiene un nodo di importazione TM1, i dettagli della connessione vengono visualizzati in questo punto.

Nome nodo. Il nome del nodo di importazione TM1.

Host di gestione. L'host di gestione Cognos TM1.

Server TM1. Il nome del server Cognos TM1.

Cubo. Il cubo TM1 da cui verranno importati i dati.

Vista. La vista da importare dal cubo TM1. È possibile importare una vista pubblica o privata.

Righe. Le righe da importare dal cubo TM1.

Colonne. Le colonne da importare dal cubo TM1.

Credenziale. Credenziale del server Cognos TM1.

Proprietà del lavoro IBM SPSS Modeler - Esportazione TM1

Se il flusso contiene un nodo di esportazione TM1, i dettagli della connessione vengono visualizzati in questo punto.

Nome nodo. Il nome del nodo di esportazione TM1.

Host di gestione. L'host di gestione Cognos TM1.

Server TM1. Il nome del server Cognos TM1.

Cubo. Il cubo TM1 in cui verranno esportati i dati.

Misura. La misura che verrà esportata.

Credenziale. Credenziale del server Cognos TM1.

Proprietà del lavoro IBM SPSS Modeler - Importazione di Analytic Server

Se il flusso contiene un nodo Origine Analytic Server, in questo punto vengono visualizzati i dettagli della connessione.

Utilizza Analytic Server predefinito. I valori sono **True** o **False**; specifica se la connessione Analytic Server definita dall'amministratore nel file `options.cfg` viene utilizzata (**True**) o se viene utilizzato un Analytic Server differente – quello specificato nel flusso/lavoro (**False**).

URL. L'URL di Analytic Server nel formato `https://hostname:port/contextroot`, dove `hostname` è l'indirizzo IP o il nome host di Analytic Server, `port` è il relativo numero di porta e `contextroot` è la root di contesto di Analytic Server.

Locatario. Il nome del locatario di cui SPSS Modeler Server è membro.

Credenziale. La credenziale per l'accesso a Analytic Server.

Nome principal del servizio. Il nome principal del servizio Kerberos.

Percorso file di configurazione. Il percorso del file di configurazione del servizio Kerberos.

Origine dati. Il nome dell'origine dati su Analytic Server.

Proprietà del lavoro IBM SPSS Modeler - Esportazione di Analytic Server

Se il flusso contiene un nodo di esportazione Analytic Server, in questo punto vengono visualizzati i dettagli della connessione.

Utilizza Analytic Server predefinito. I valori sono **True** o **False**; specifica se la connessione Analytic Server definita dall'amministratore nel file `options.cfg` viene utilizzata (**True**) o se viene utilizzato un Analytic Server differente – quello specificato nel flusso/lavoro (**False**).

URL. L'URL di Analytic Server nel formato `https://hostname:port/contextroot`, dove `hostname` è l'indirizzo IP o il nome host di Analytic Server, `port` è il relativo numero di porta e `contextroot` è la root di contesto di Analytic Server.

Locatario. Il nome del locatario di cui SPSS Modeler Server è membro.

Credenziale. La credenziale per l'accesso a Analytic Server.

Nome principal del servizio. Il nome principal del servizio Kerberos.

Percorso file di configurazione. Il percorso del file di configurazione del servizio Kerberos.

Origine dati. Il nome dell'origine dati su Analytic Server.

IBM SPSS Modeler Proprietà del lavoro - Notifiche

Consente di specificare le notifiche email per gli errori o la riuscita della fase di lavoro.

Fare clic sul pulsante **Aggiorna** in ciascun caso per aggiungere o eliminare i destinatari delle notifiche.

Visualizzazione dei flussi IBM SPSS Modeler

È possibile avviare l'applicazione IBM SPSS Modeler direttamente da IBM SPSS Deployment Manager.

Per visualizzare il flusso nell'applicazione IBM SPSS Modeler, fare doppio clic sul flusso in Esplora contenuto. Il sistema avvia l'applicazione IBM SPSS Modeler e visualizza il flusso in IBM SPSS Modeler.

E' importante tenere presente che se vengono apportate modifiche a un file in un lavoro, ad esempio, a un file di flusso IBM SPSS Modeler (.str)--tutti i lavori che contengono file sono interessati dalla modifica. Quando si effettuano modifiche in un file, viene salvata una nuova versione del file nel repository. Tuttavia, il lavoro che contiene il file non viene automaticamente aggiornato con il file modificato. Per incorporare gli aggiornamenti del file nel lavoro interessato:

1. Aprire di nuovo il lavoro. Quando il lavoro viene riaperto, viene visualizzato un asterisco con il nome lavoro nell'area di lavoro per indicare che il lavoro contiene modifiche non salvate.
2. Salvare di nuovo il lavoro.

Codici di completamento di IBM SPSS Modeler

I codici di completamento dei lavori di IBM SPSS Modeler sono descritti nella tabella seguente. Utilizzare questi codici di completamento per ciascuna relazione condizionale che coinvolge i flussi di IBM SPSS Modeler.

Tabella 1. Codici di completamento dei lavori di IBM SPSS Modeler

Codice	Descrizione
0	operazione completata
1	errore nell'esecuzione del flusso
2	errore nella pubblicazione
8	errore sconosciuto

Limitazioni dei flussi di IBM SPSS Modeler

Quando si utilizzano i flussi in IBM SPSS Deployment Manager, il sistema presenta le seguenti limitazioni:

- **Denominazione.** Se il nome, l'etichetta e il tipo del nodo sono uguali, non è possibile programmare l'esecuzione di un lavoro, perché viene generato un conflitto.
- **Script.** Se uno script del flusso contiene sovrascritture per un nodo specifico ad esempio, lo script imposta la posizione di output per un grafico, sostituisce qualsiasi valore in conflitto specificato dall'utente definito in Deployment Manager.
- **Supernodi.** L'esecuzione della fase di un lavoro di IBM SPSS Modeler corrisponde all'elaborazione ed esecuzione di tutti i nodi terminali di primo livello nel flusso. Se uno dei nodi terminali è un supernodo terminale, i nodi terminali all'interno di questo supernodo vengono eseguiti in modo ricorsivo. Al contrario, se i supernodi sorgente o processo, che sono per definizione non terminali, contengono nodi terminali, tali nodi non vengono eseguiti. I nodi terminali nei supernodi non terminali vengono visualizzati in Deployment Manager ma non vengono elaborati durante l'esecuzione della fase.
- **Parametri.** I parametri definiti per i supernodi non possono avere lo stesso nome dei parametri definiti per un flusso. I nomi devono essere univoci.
- **Limitazioni dei nodi Text Analytics.** Non è possibile utilizzare i nodi elenco file o feed web di SPSS Modeler Text Analytics per il calcolo del punteggio all'interno di una configurazione IBM SPSS Collaboration and Deployment Services - Scoring.

Oltre agli elementi sopra riportati, esistono i vincoli specifici riportati di seguito.

- Quando si archivia un flusso nel repository IBM SPSS Collaboration and Deployment Services, non è disponibile alcuna casella di controllo di sblocco. Lo sblocco è il funzionamento predefinito durante l'archiviazione dei flussi. Per bloccare o sbloccare un oggetto, selezionare **Strumenti > Repository > Esplora**, passare all'oggetto e fare clic con il tasto destro del mouse sul relativo nome per visualizzare il menu di scelta rapida.
- In Deployment Manager, quando viene eseguito un lavoro che contiene un flusso con un nodo Valutazione impostato in modo da creare un grafico dei profitti, l'output del grafico potrebbe non essere completo se il sistema è in esecuzione in Oracle Weblogic 11g utilizzando Oracle JRockit JRE. Per evitare questo problema, utilizzare IBM JRE.

Tipi di nodi

Quando si apre un flusso IBM SPSS Modeler da IBM SPSS Deployment Manager, si può notare che i nodi del flusso sono rappresentati mediante icone di forma diversa. Le icone circolari rappresentano i nodi di dati di input, mentre quelle esagonali rappresentano operazioni di processi eseguiti su record e campi di dati. I triangoli indicano un output grafico e i pentagoni rappresentano i nodi di modellazione. L'output (non grafico) e le operazioni di esportazione sono indicati da nodi rettangolari.

I diversi nodi sono pienamente descritti nelle guide *IBM SPSS Modeler Nodi origine, di elaborazione ed output* e nel nodo modellazione *IBM SPSS Modeler*.

Esempio script: Confronto tra i modelli Campione e Sfidante

Il seguente estratto di uno script di IBM SPSS Modeler mostra un confronto campione/sfidante. Il modello campione viene confrontato con il modello sfidante. Se il modello sfidante è migliore del modello campione, il modello campione viene sovrascritto. Pertanto, il modello sfidante diventa il nuovo modello campione.

```
# -----
# Before rebuilding model, examine performance of existing model
# -----
#--- LOAD Model
set working_file = ^SPSS_SMS_INT_PERSIST_PATH >> ^slash >> "ResolutionStatus.gm"
load model ^working_file

#--- First do the evaluation of the current champion model
delete 'ResolutionStatusModel'
insert model 'ResolutionStatus' at 558 652
rename 'ResolutionStatus':applyc50node as 'ResolutionStatusModel'
connect 'ValidationSet' to 'ResolutionStatusModel'
connect 'ResolutionStatusModel' to 'PredictQuality'

set 'ModelEval'.full_filename = ^SPSS_SMS_INT_TEMP_PATH >> ^slash >> "ModelEvalChamp.txt"
execute 'ModelEval':tablenode
var missRateChamp
set missRateChamp = value 'ModelEval':tablenode.output at 1 6
delete 'ResolutionStatusModel'
clear generated palette

#--- Execute c50node 'ResStatusModelBuild'
if ^useBoost = 1
  set 'ResStatusModelBuild'.use_boost = T
  set 'ResStatusModelBuild'.boost_num_trials = ^numBoostTrials
else
  set 'ResStatusModelBuild'.use_boost = F
endif
if ^useMisCosts = 1
  set 'ResStatusModelBuild'.use_costs = T
  set 'ResStatusModelBuild'.costs = [{'Not fraud' 'Not fraud' 0}{'Not fraud' 'Fraud' ^falsePosCost}
{'Fraud' 'Not fraud' ^falseNegCost}{'Fraud' 'Fraud' 0}]
endif
execute 'ResStatusModelBuild'
```

```

#--- Now evaluate the new model
insert model 'ResolutionStatus' at 558 652
rename 'ResolutionStatus':applyc50node as 'ResolutionStatusModel'
connect 'ValidationSet' to 'ResolutionStatusModel'
connect 'ResolutionStatusModel' to 'PredictQuality'

set 'ModelEval'.full_filename = ^SPSS_SMS_INT_TEMP_PATH >> ^slash >> "ModelEvalChallenge.txt"
execute 'ModelEval':tablenode
var missRateChallenge
set missRateChallenge = value 'ModelEval':tablenode.output at 1 6

#--- Determine which one should be the champion
#--- Overwrite existing champion if challenger is better

if ^missRateChamp > ^missRateChallenge
#--- SAVE Model
set working_file = ^SPSS_SMS_INT_PERSIST_PATH >> ^slash >> "ResolutionStatus.gm"
save model 'ResolutionStatus':applyc50node as ^working_file

#--- EXPORT Model
set working_file = ^SPSS_SMS_INT_PERSIST_PATH >> ^slash >> "ResolutionStatus.xml"
export model 'ResolutionStatus' as ^working_file format pmm1

#--- EXPORT Model
set working_file = ^SPSS_SMS_INT_PERSIST_PATH >> ^slash >> "ResolutionStatus_summary.html"
export model 'ResolutionStatus' summary as ^working_file format html
endif

```

Capitolo 4. Servizio di calcolo del punteggio

Il Servizio di calcolo del punteggio consente alle applicazioni client di utilizzare i punteggi in tempo reale derivati dai modelli predittivi sviluppati in IBM SPSS Modeler. Il servizio richiama il modello specificato, lo carica, richiama l'implementazione del calcolo del punteggio corretta e restituisce il risultato al client.

Il calcolo del punteggio è il processo di generazione di valori in tempo reale mediante la fornitura di dati di input ai modelli predittivi. Un modello di calcolo del punteggio è una qualsiasi risorsa utente che può essere utilizzata per produrre dei valori di output, forniti dei dati di input. In generale, per utilizzare un modello per la generazione di punteggi:

1. Selezionare un modello da utilizzare per il calcolo del punteggio da IBM SPSS Collaboration and Deployment Services Repository.
2. Definire una configurazione di calcolo del punteggio per il modello.
3. Fornire i dati al modello configurato e generare i punteggi.

Per ulteriori informazioni relative al calcolo del punteggio ed al Servizio di calcolo del punteggio, consultare la documentazione di IBM SPSS Collaboration and Deployment Services.

Limitazioni dei flussi di IBM SPSS Modeler

Quando si utilizzano i flussi in IBM SPSS Deployment Manager, il sistema presenta le seguenti limitazioni per il Servizio di calcolo del punteggio:

- **Supernodi.** I nodi di origine ed i nodi terminali all'interno di supernodi non sono supportati
- **Nodi geospaziali.** I nodi geospaziali non sono supportati.
- **Nodi builder del modello.** Il nodo builder del modello come nodo terminale non è supportato.
- **Mining nel database.** I flussi che contengono i nugget di mining in-database non sono supportati.
- **UDF, UDA e WUDA.** I flussi che utilizzano UDF, UDA e WUDA non sono supportati.
- **Nodi origine** I nodi di origine con un modello dati di output, contenente un tipo elenco, non sono supportati.
- **Nodi terminali.** I nodi terminali con un modello di dati di input, contenente un tipo elenco, non sono supportati.
- **Nodi Estensione Python for Spark.** I flussi che contengono i seguenti nodi non sono supportati: nodo Importazione estensione - Python for Spark, nodo Trasformazione estensione - Python for Spark, nodo Modello di estensione - Python for Spark, Nugget del modello Estensione Python, Output estensione - Python for Spark, Esportazione estensione - Python for Spark.
- **Limitazioni dei nodi Text Analytics.** Non è possibile utilizzare i nodi elenco file o feed web di SPSS Modeler Text Analytics per il calcolo del punteggio all'interno di una configurazione IBM SPSS Collaboration and Deployment Services - Scoring.
- **Nodi sorgente dell'Analytic Server.** Se si dispone di un flusso SPSS Modeler che utilizza il nodo sorgente di Analytic Server, è necessario consentire una connessione diretta tra il IBM SPSS Collaboration and Deployment Services Server e l'Analytic Server. Altrimenti, il job fallirà se il firewall bloccherà la connessione tra i due server.

Capitolo 5. Fasi di lavoro campione/sfidante

Panoramica campione/sfidante

Utilizzando IBM SPSS Deployment Manager, è possibile confrontare i file di modello generati da IBM SPSS Modeler per determinare quale file contiene il modello predittivo più efficace. La fase di lavoro campione/sfidante valuta un modello confrontandolo con uno o più sfidanti.

Una volta eseguito il confronto dei risultati, il modello migliore diventa il nuovo campione.

Campione. Il campione corrisponde al modello più efficace. Per l'esecuzione iniziale della fase di lavoro campione/sfidante, non è disponibile alcun campione, ma soltanto il primo sfidante e l'elenco degli sfidanti corrispondente. Per le successive esecuzioni della fase di lavoro, il sistema determinerà il campione.

Sfidante. Gli sfidanti vengono confrontati tra loro. Lo sfidante che genera i migliori risultati diventa il nuovo campione.

Processo di selezione del campione

Il processo di confronto campione/sfidante prevede le seguenti attività:

1. Calcolo del punteggio di ogni modello in competizione.
2. Valutazione dei risultati del calcolo del punteggio.
3. Confronto dei risultati di valutazione e individuazione dello sfidante campione.
4. Salvataggio del nuovo campione nel repository (facoltativa).

Aggiunta di una fase di confronto campione/sfidante a un lavoro

Per aggiungere una fase lavoro campione/sfidante ad un lavoro, selezionare lo strumento Campione/sfidante dalla palette lavoro e fare clic nell'area di lavoro.

Metrica di valutazione dei modelli

Il confronto e la valutazione del modello possono incentrarsi sulla precisione, profitti o accreditamento.

- **Accuratezza.** La precisione di un modello riflette la percentuale di risposte all'obiettivo che si prevede correttamente. I modelli aventi un'alta percentuale di previsioni corrette vengono preferiti a quelli aventi una bassa percentuale.
- **Profitti** La statistica dei profitti è un indicatore delle prestazioni di un modello. Questa misura confronta i risultati da un modello rispetto ai risultati ottenuti senza utilizzare un modello. Il miglioramento nei risultati ottenuti quando si utilizza un modello fa riferimento ai profitti. Quando si confrontano due modelli, è preferito il modello avente il valore di profitti più alto rispetto al percentile specificato
- **Accreditamento.** L'accREDITamento del modello riflette l'attendibilità del modello. Questo approccio esamina le similitudini tra i nuovi dati e i dati di addestramento su cui si basa un modello. I valori di accreditamento variano da 0 a 1 con i valori più alti che indicano una maggiore similitudine tra i predittori nei due insiemi di dati. Quando si confrontano due modelli, il modello avente il valore di accreditamento più alto si basa sui dati di addestramento che sono più simili ai nuovi dati, rendendo quindi il modello più attendibile e preferito.

Dipendenza dall'ordine

A differenza di tutti gli altri tipi di fasi lavoro, le schede della fase di lavoro campione/sfidante dipendono dall'ordine.

Ad esempio, è necessario selezionare un'opzione nella scheda Sfidanti per poter modificare le informazioni nella scheda Campione. Inoltre, le informazioni visualizzate in alcune tabelle dipendono dagli sfidanti selezionati nella tabella Sfidanti.

Il processo di esecuzione di un confronto campione/sfidante è costituito dalle seguenti fasi:

1. Immissione di informazioni generali sul lavoro.
2. Identificazione degli sfidanti.
3. Specificazione delle informazioni relative al campione.
4. Visualizzazione delle informazioni dei parametri.
5. Specificazione delle notifiche.

Informazioni generali

La scheda generale contiene le informazioni che riguardano la fase lavoro generale del campione.

Sono richieste le informazioni relative al server e all'accesso IBM SPSS Modeler sono richieste per eseguire un fase lavoro campione/sfidante. I dati di accesso e del server content repository sono necessari per eseguire il lavoro e salvare i nuovi campioni nel IBM SPSS Collaboration and Deployment Services Repository (i dati di accesso e del server content repository sono necessari anche se si scartano i risultati dell'analisi). Le credenziali sono basate sull'utente attualmente collegato al sistema.

Nome fase lavoro. Il nome della fase di lavoro. Per impostazione predefinita, il nome della prima fase lavoro è *Evento 1*. Le successive fasi lavoro vengono denominate *Evento 2*, *Evento 3* e così via. Il nome specificato qui appare nella tabella della cronologia dei lavori dopo l'esecuzione della fase di lavoro.

Server IBM SPSS Modeler. Server o cluster di server IBM SPSS Modeler su cui verrà eseguito il flusso. L'elenco contiene tutti i server e i cluster di server attualmente configurati per eseguire fasi di IBM SPSS Modeler. Per modificare il server, selezionarlo dall'elenco a discesa **IBM SPSS Modeler Server**. Per creare una nuova definizione server, fare clic su **Nuovo** per avviare la procedura guidata di definizione server.

IBM SPSS Modeleraccesso. Credenziali utilizzate per accedere al server o al cluster di server di IBM SPSS Modeler. Per modificare le credenziali, selezionare una definizione di credenziale dall'elenco a discesa **IBM SPSS Modeler Login**. Per definire nuove credenziali, fare clic su **Nuovo** per avviare la procedura guidata di definizione delle credenziali.

Content Repository Server. Il Content Repository Server consente a un lavoro di salvare i file in un IBM SPSS Collaboration and Deployment Services Repository. Generalmente, il Content Repository Server viene specificato quando si aggiornano i modelli mediante IBM SPSS Modeler. Per specificare un Content Repository Server, selezionare un server dall'elenco a discesa **Content Repository Server**. Per creare una nuova definizione server, fare clic su **Nuovo** per avviare la procedura guidata di definizione server. Per generare una definizione Content Repository Server in base alle informazioni server correnti, fare clic su **Genera**. Viene creata e popolata automaticamente una definizione server nel campo *Content Repository Server*.

Accesso a repository di contenuti Informazioni di accesso per il Content Repository Server. Per specificare le informazioni di accesso al repository di contenuti, selezionare una credenziale dall'elenco a discesa **Accesso a repository di contenuti**. Per creare un nuovo accesso, fare clic su **Nuovo** per avviare la procedura guidata di accesso al repository di contenuti. Se non viene utilizzata la funzione Single sign-on (SSO) per la connessione al IBM SPSS Collaboration and Deployment Services Repository, fare clic su **Genera** per generare un accesso per il Content Repository Server basato sulle impostazioni di protezione

esistenti. Viene creato e popolato automaticamente un accesso per il repository di contenuti nel campo *Accesso a repository di contenuti*. La generazione dell'accesso non è disponibile quando si utilizza la funzione Single sign-on (SSO).

Sfidanti

È necessario selezionare almeno un primo sfidante per eseguire una fase di lavoro campione/sfidante. È importante notare che il primo sfidante selezionato non implica un ordine di visualizzazione né alcuna priorità nel processo di valutazione. Il primo sfidante è semplicemente un riferimento.

Sorgente dati ed etichette utilizzate per determinare gli sfidanti successivi vengono stabilite in base al primo sfidante. Una volta selezionato o aggiornato il primo sfidante, i campi rimanenti nella tabella vengono aggiornati con le informazioni corrispondenti al primo sfidante.

Primo sfidante. Nome del primo sfidante. Per selezionarlo dal repository, fare clic su **Sfoglia**.

Etichetta primo sfidante. Etichetta associata al file di modello contenente il primo sfidante. Specificare questo valore quando si seleziona il primo sfidante.

Sfidante origine dati. Lo sfidante che fornisce il nodo della sorgente dati utilizzato per la fase di lavoro. Fare clic su **Sfoglia** per scegliere questo sfidante nell'elenco di voci selezionate nella casella Sfidanti.

Metrica. Criterio di misurazione in base al quale vengono confrontati gli sfidanti. I valori validi comprendono *accreditamento*, *precisione e profitti*. Se si seleziona *Guadagni*, è necessario specificare anche un percentile. Consultare l'argomento "Metrica di valutazione dei modelli" a pagina 25 per ulteriori informazioni.

Tabella Sfidanti

La tabella Sfidanti elenca i rami dei punteggi di default per gli sfidanti che corrispondono alla sorgente dati e l'etichetta associata al primo sfidante. Solo gli sfidanti selezionati dalla tabella verranno confrontati con il primo sfidante durante l'esecuzione della fase di lavoro. Selezionando (o deselegando) uno sfidante dall'elenco, il sistema di conseguenza aggiornerà le corrispondenti informazioni nelle altre schede della fase lavoro.

Ogni volta che viene aperto un lavoro contenente una fase campione/sfidante salvata in precedenza, l'elenco degli sfidanti viene automaticamente aggiornato. I nuovi sfidanti possono essere aggiunti all'elenco se soddisfano i criteri di sorgente dati ed etichetta del primo sfidante. Al contrario, gli sfidanti che non soddisfano più tali criteri possono essere rimossi dall'elenco degli sfidanti. Se uno sfidante selezionato è stato rimosso dal repository, il sistema genera un messaggio per indicare che lo sfidante non è più disponibile.

Sebbene gli sfidanti possono essere selezionati e deselegati per effettuare il confronto, le informazioni contenute nella tabella Sfidanti non possono essere modificate. In particolare, la tabella Sfidanti contiene le informazioni riportate di seguito.

Nome. Nome dello sfidante.

Etichetta. Etichetta associata allo sfidante.

Descrizione Descrizione dello sfidante.

Modifiche al primo sfidante

Le modifiche apportate al primo sfidante dopo il salvataggio del lavoro possono influire sull'analisi dello sfidante campione. Ad esempio, si supponga che il primo sfidante venga rimosso dal repository o che

venga rimossa l'etichetta associata al primo sfidante. Quando si accede alla scheda Sfidanti, il sistema genera un messaggio per indicare che il primo sfidante non è più disponibile per l'uso. In questo caso, è necessario specificare un nuovo primo sfidante.

Selezione degli sfidanti

Per selezionare i modelli sfidante da includere nelle analisi campione/sfidante, attenersi alla procedura descritta di seguito:

1. Nella scheda Sfidanti di una fase Campione/Sfidante, fare clic su **Sfoglia** per il Primo sfidante. Per aggiungere manualmente gli sfidanti, fare clic su **Aggiungi** per la tabella Sfidanti.
2. Selezionare il file di modello facendo clic su **Sfoglia**. Il file di modello è un flusso IBM SPSS Modeler contenente un ramo del punteggio di default con un nugget del modello valido.
3. Selezionare l'etichetta che indica la versione del file di modello da utilizzare.
4. Nella tabella Sfidanti, selezionare il ramo del punteggio da utilizzare.
5. Fare clic su **OK**.

Sfidanti non validi

Per poter essere confrontati i file di modello devono avere rami di calcolo del punteggio che utilizzano una struttura dati comune.

Le caratteristiche dei dati che devono coincidere per tutti gli sfidanti includono:

- Le sorgenti dati devono presentare il medesimo numero di campi.
- I nomi dei campi devono essere identici in tutte le sorgenti dati.
- I livelli di misurazione dei campi devono essere identici in tutte le sorgenti dati.

Se il sistema non è in grado di trovare sfidanti compatibili con il primo sfidante selezionato, viene visualizzata la finestra di dialogo Sfidante non valido. Per selezionare un nuovo sfidante:

1. Fare clic su **OK** per tornare alla scheda Sfidanti.
2. Selezionare un nuovo sfidante.

Selezione delle sorgenti dati degli sfidanti

Per selezionare le sorgenti dati da utilizzare nelle analisi campione/sfidante, attenersi alla procedura descritta di seguito:

1. Nella scheda Sfidanti di una fase Campione/sfidante, fare clic su **Sfoglia** per lo sfidante origine dati.
2. Dall'elenco degli sfidanti inclusi nell'analisi, selezionare il ramo del punteggio che include la sorgente dati da utilizzare.
3. Fare clic su **OK**.

Campione

Affinché sia possibile specificare le informazioni per un campione, è necessario selezionare almeno uno sfidante. Se si accede alla scheda Campione prima di selezionare uno sfidante, viene visualizzata una finestra di dialogo per indicare che il primo sfidante non è stato selezionato e che è necessario selezionarlo.

Non creare una nuova versione del campione. Selezionare questa opzione per evitare la creazione di una nuova versione del campione. In questo caso, la versione etichettata selezionata del campione verrà modificata. Deselezionare questa opzione per creare una nuova versione del campione invece di modificare quella etichettata.

Nome file. Nome da utilizzare per la copia dello sfidante identificata come campione.

Posizione. Percorso in cui è archiviata la copia del file campione.

Autorizzazione. Autorizzazioni associate alla copia del campione.

Metadati. Proprietà associate alla copia del campione. Specificare i metadati per l'output del campione è come specificare i metadati per altri tipi di output del lavoro.

Utilizzo del campione in altri lavori

Una volta eseguito il lavoro campione/sfidante, il campione ottenuto può essere utilizzato in altri lavori. Per includere il campione in un altro lavoro, sono necessarie le seguenti informazioni:

- Nome del campione.
- Percorso del campione.

Quando il campione viene utilizzato in un altro lavoro, viene applicata l'etichetta *LATEST*. Questa etichetta non può essere modificata.

Prova del campione

Di default, il sistema crea una nuova copia del campione ogni volta che si esegue il lavoro campione/sfidante, archivia la copia nel percorso di output specificato e scrive i risultati nel registro della cronologia del lavoro.

Tuttavia, possono esserci istanze in cui non si desidera salvare una copia del campione. Ad esempio, si supponga di voler semplicemente testare la fase di lavoro campione/sfidante.

Per disabilitare la creazione di una copia, selezionare la casella di controllo **Non creare una nuova versione del campione**. Se questa casella di controllo è selezionata, le opzioni rimanenti della scheda non sono disponibili. Il sistema utilizzerà le stesse informazioni applicate al campione corrente.

Il sistema esegue il confronto campione/sfidante e determina un nuovo campione. Tuttavia, una nuova versione del campione non verrà creata né salvata nel repository. I risultati verranno invece scritti nel registro della cronologia del lavoro e verrà indicato lo sfidante scelto come campione.

Ad esempio, supponiamo che sia stato eseguito un lavoro campione/sfidante e che il creatore del lavoro abbia scelto di non creare una nuova versione del campione. Il registro della cronologia del lavoro risultante potrebbe avere il seguente aspetto:

```
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.22 sec, CPU=18.97 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
```

```

Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.2 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.06 sec, CPU=17.75 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.48 sec, CPU=0.19 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=21.48 sec, CPU=17.34 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.17 sec
The result for challenger cc_cartresponse.str is 98.809.
The result for challenger cc_neuralnetresponse.str is 98.844.
The result for challenger cc_c51response.str is 98.809.
The declared Champion is cc_neuralnetresponse.str.

```

Notare l'ultima riga nel file di log:

```
The declared Champion is cc_neuralnetresponse.str.
```

Questa riga indica che il flusso *cc_neuralnetresponse* dovrebbe essere stato il campione. Tuttavia, una copia di questo flusso non è stato salvato nel repository perché il sistema non ha creato una nuova versione del flusso. Se il flusso è stato salvato nel repository, il log dovrebbe contenere una riga aggiuntiva che indica che il flusso è stato salvato nel repository, ad esempio:

```
Adding artifact spsstr:/PMDemo/ModelManagement/cc_neuralnetresponse.str.
```

File di dati

Le informazioni relative ai file di dati appaiono per gli sfidanti selezionati dall'elenco di sfidanti nella scheda Sfidanti.

Eventuali modifiche apportate in questa scheda sono valide solo per la fase di lavoro campione/sfidante. Le modifiche apportate alle informazioni dei file di dati non vengono propagate allo sfidante salvato nel IBM SPSS Collaboration and Deployment Services Repository. La tabella dei file di dati contiene le seguenti informazioni.

Nome nodo. Nome del nodo origine che contiene i dati utilizzati dal flusso. Il nome del nodo non è modificabile.

Tipo di nodo. Tipo di nodo definito nel flusso. Il tipo di nodo non è modificabile.

Nome file. Nome del file di dati di input. Per modificare il nome, fare clic nella cella del nome file e modificare il nome.

Formato. Formato del file di output, per esempio un file delimitato da virgole. Per modificare il tipo di formato del file, fare clic nella cella Formato. Viene visualizzato un elenco a discesa. Selezionare il tipo di formato.

Posizione. Posizione dei file di dati di input. Per modificare la posizione, fare clic nella colonna e quindi sul pulsante con i puntini di sospensione così visualizzato. Viene visualizzata la finestra di dialogo Posizione file di input. Modificare la posizione in funzione delle proprie esigenze.

Vista dati

Le informazioni relative alla vista dati analitici appaiono per gli sfidanti selezionati dall'elenco di sfidanti nella scheda Sfidanti.

Eventuali modifiche apportate in questa scheda sono valide solo per la fase di lavoro campione/sfidante. Le modifiche apportate alle informazioni della vista dati non vengono propagate allo sfidante salvato nel IBM SPSS Collaboration and Deployment Services Repository. La tabella della vista dati contiene le seguenti informazioni.

Nome nodo. Nome del nodo Visualizzazione dati che contiene i dati utilizzati dal flusso. Il nome del nodo non è modificabile.

Vista dati analitici. La vista dati analitici a cui fa riferimento il nodo Visualizzazione dati.

Etichetta. L'etichetta che identifica la versione della vista dati analitici utilizzata.

Nome tabella. La tabella contenente i campi dei dati di input.

Piano di accesso ai dati. Il piano che fornisce i record di dati per i campi dei dati di input. Per modificare il piano di accesso ai dati utilizzato per un nodo, selezionare la cella che contiene il piano di accesso ai dati e fare clic sul pulsante con i puntini di sospensione (...) così visualizzato.

origine dati ODBC

Le informazioni relative alle sorgenti dati ODBC appaiono per gli sfidanti selezionati dall'elenco di sfidanti nella scheda Sfidanti.

Eventuali modifiche apportate in questa scheda sono valide solo per la fase di lavoro campione/sfidante. Le modifiche apportate alle informazioni sulle sorgenti dati ODBC non vengono propagate all'indietro nello sfidante nel IBM SPSS Collaboration and Deployment Services Repository. La tabella Origini dati ODBC contiene le seguenti informazioni.

Nome nodo. Nome del nodo origine che contiene i dati utilizzati dal flusso. Il nome è preceduto da un prefisso composto dal nome di ogni supernodo contenente il nodo, racchiuso tra barre. Ad esempio, se il nodo *MioNodo* è contenuto in un supernodo denominato *Supernodo1*, il nome viene visualizzato in questo modo */Supernodo1/MioNodo*.

Tipo di nodo. Tipo di nodo definito nel flusso.

Origini dati ODBC. Il nome corrente della sorgente dati ODBC (DSN). Per passare ad una sorgente dati ODBC diversa, fare clic sulla cella contenente il nome della sorgente dati corrente, quindi fare clic sul "..." pulsante che viene visualizzato. In questo modo viene visualizzata una finestra di dialogo da dove è possibile scegliere un DSN esistente o crearne uno nuovo. Notare che le impostazioni della fase lavoro sovrascrivono sempre il flusso; per se si modifica il flusso per utilizzare una diversa origini dati, è necessario anche modificare il lavoro per utilizzare la stessa origine, altrimenti l'esecuzione del lavoro può avere esito negativo.

Credenziali. Per modificare il nome utente e la password del database quando si cambia la sorgente dati ODBC, fare clic sulla cella contenente le credenziali correnti, quindi fare clic sul "..." pulsante che viene visualizzato. In questo modo viene visualizzata una finestra di dialogo da dove è possibile scegliere una definizione di credenziali esistente o crearne una nuova.

Tabella di database. Tabella di database che corrisponde al nodo.

I nodi all'interno di supernodi bloccati non sono accessibili. Non possono essere né visualizzati né modificati.

Importazione Cognos

Se i file di modello contengono nodi origine IBM Cognos BI, vengono visualizzati i dettagli della connessione Cognos.

Nome nodo. Nome del nodo origine Cognos.

URL di connessione. URL del server Cognos al quale viene stabilita la connessione.

Nome package. Nome del pacchetto Cognos da cui vengono importati i metadati.

Anonimo. Include **Anonymous** se viene effettuato un accesso anonimo per la connessione al server Cognos oppure **Credential** se vengono utilizzati un nome utente e una password Cognos.

Credenziali. Il nome utente e la password (se richiesta) sul server Cognos.

Nota: Le credenziali Cognos devono essere create in un dominio che rappresenta l'ID dello spazio dei nomi Cognos.

Informazioni particolari

Queste informazioni sono state sviluppate per prodotti e servizi offerti negli Stati Uniti. Questo materiale potrebbe essere disponibile da IBM in altre lingue. Tuttavia, potrebbe essere necessario disporre di una propria copia del prodotto o versione di prodotto in quella lingua per potervi accedere.

IBM può non offrire i prodotti, i servizi o le funzioni presentati in questo documento in altri paesi. Consultare il rappresentante locale IBM per le informazioni sui prodotti e servizi attualmente disponibili nella propria zona. Qualsiasi riferimento ad un prodotto, programma o servizio IBM non implica o intende dichiarare che solo quel prodotto, programma o servizio IBM può essere utilizzato. In sostituzione a quelli forniti da IBM, è possibile utilizzare prodotti, programmi o servizi funzionalmente equivalenti che non comportino violazione dei diritti di proprietà intellettuale o di altri diritti IBM. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La consegna del presente documento non conferisce alcuna licenza rispetto a questi brevetti. Chi desiderasse ricevere informazioni relative a licenze può rivolgersi per iscritto a:

*IBM Director of Licensing
IBM Europe
North Castle Drive, MD-NC119
D-7030 Boeblingen
US*

Per richieste di licenze relative ad informazioni double-byte (DBCS) contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION) FORNISCE LA PRESENTE PUBBLICAZIONE "NELLO STATO IN CUI SI TROVA" SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, IVI INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI NON VIOLAZIONE, DI COMMERCIALIZZABILITÀ E DI IDONEITÀ PER UNO SCOPO PARTICOLARE. Alcune giurisdizioni non escludono le garanzie implicite; di conseguenza la suddetta esclusione potrebbe, in questo caso, non essere applicabile.

Le presenti informazioni possono includere imprecisioni tecniche o errori tipografici. Le modifiche periodiche apportate alle informazioni contenute in questa pubblicazione verranno inserite nelle nuove edizioni della pubblicazione. IBM si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o al programma descritto nel manuale in qualsiasi momento e senza preavviso.

Tutti i riferimenti a siti Web non IBM sono forniti unicamente a scopo di consultazione e non devono essere in alcun modo considerati come complementari a tali siti Web. I materiali disponibili su tali siti Web non fanno parte del materiale relativo a questo prodotto IBM e l'utilizzo di questi è a discrezione dell'utente.

IBM può utilizzare o distribuire qualsiasi informazione fornita dall'utente nel modo che ritiene più idoneo senza incorrere in alcun obbligo nei confronti dell'utente stesso.

Coloro che detengono la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) lo scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'uso reciproco di tali informazioni dovrebbero contattare:

IBM Director of Licensing
IBM Europe
North Castle Drive, MD-NC119
D-7030 Boeblingen
US

Queste informazioni possono essere rese disponibili secondo condizioni contrattuali appropriate, compreso, in alcuni casi, l'addebito di un canone.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale concesso in licenza disponibile sono forniti da IBM in base ai termini dell'IBM Customer Agreement, dell'IBM International Program License Agreement o di qualsiasi altro accordo equivalente tra le parti.

I dati delle prestazioni e gli esempi client citati vengono presentati solo a scopo illustrativo. I risultati delle prestazioni effettive possono variare in base alle configurazioni specifiche e alle condizioni di funzionamento.

Le informazioni relative a prodotti non IBM sono ottenute dai fornitori di quei prodotti, dagli annunci pubblicati o da altre fonti disponibili al pubblico. IBM non ha testato quei prodotti e non può garantire l'accuratezza delle prestazioni, la compatibilità o qualsiasi altra dichiarazione relativa a prodotti non IBM. Commenti relativi alle prestazioni di prodotti non IBM, dovrebbero essere indirizzati ai fornitori di questi prodotti.

Qualsiasi affermazione relativa agli obiettivi e alla direzione futura di IBM è soggetta a modifica o revoca senza preavviso e concerne esclusivamente gli scopi dell'azienda.

Questa pubblicazione contiene esempi di dati e prospetti utilizzati quotidianamente nelle operazioni aziendali. Per fornire una descrizione il più possibile esaustiva, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti questi nomi sono fittizi e qualsiasi somiglianza a persone o aziende commerciali reali è puramente casuale.

Marchi

IBM, il logo IBM e ibm.com sono marchi o marchi registrati di International Business Machines Corp., registrati in numerose giurisdizioni del mondo. I nomi di altri prodotti e servizi potrebbero essere marchi di IBM o di altre società. Per un elenco aggiornato di marchi IBM, consultare il web nella sezione Copyright and trademark information, all'indirizzo www.ibm.com/legal/copytrade.shtml.

Adobe, il logo Adobe logo, PostScript ed il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o relative controllate negli Stati Uniti e altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o in altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o in altri paesi.

UNIX è un marchio registrato di Open Group negli Stati Uniti e/o in altri paesi.

Java e tutti i marchi e i logo relativi a Java sono marchi commerciali o marchi registrati di Oracle e/o delle sue affiliate.

Termini e condizioni per la documentazione del prodotto

Le autorizzazioni per l'uso delle presenti pubblicazioni sono concesse in conformità con i seguenti termini e condizioni.

Applicabilità

I termini e le condizioni riportati di seguito si aggiungono alle condizioni di utilizzo per il sito Web IBM.

Uso personale

È possibile riprodurre tali pubblicazioni per uso personale e non commerciale nel rispetto di tutte le informazioni relative alla proprietà. Non è possibile distribuire, visualizzare o utilizzare tali pubblicazioni, o una parte di esse, senza l'esplicito consenso di IBM.

Uso commerciale

È possibile riprodurre, distribuire e visualizzare queste pubblicazioni unicamente all'interno del proprio gruppo aziendale a condizione che vengano conservate tutte le indicazioni relative alla proprietà. Non è possibile effettuare lavori derivati di queste pubblicazioni o riprodurre, distribuire o visualizzare queste pubblicazioni o qualsiasi loro parte al di fuori del proprio gruppo aziendale senza chiaro consenso da parte di IBM.

Diritti

Fatto salvo quanto espressamente concesso in questa autorizzazione, non sono concesse altre autorizzazioni, licenze o diritti, espressi o impliciti, relativi alle pubblicazioni o a qualsiasi informazione, dato, software o altra proprietà intellettuale qui contenuta.

IBM si riserva il diritto di ritirare le autorizzazioni qui concesse qualora, a propria discrezione, l'utilizzo di queste Pubblicazioni sia a danno dei propri interessi o, come determinato da IBM, qualora non siano rispettate in modo appropriato le suddette istruzioni.

Non è consentito scaricare, esportare o riesportare queste informazioni se non nei limiti stabiliti dalle leggi e normative applicabili, ivi comprese tutte le leggi e le normative sull'esportazione vigenti negli Stati Uniti.

IBM NON GARANTISCE IL CONTENUTO DI QUESTE PUBBLICAZIONI. ESSE SONO FORNITE "NELLO STATO IN CUI SI TROVANO", SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, INCLUSE, A TITOLO ESEMPLIFICATIVO, GARANZIE DI COMMERCIALIZZABILITÀ, IDONEITÀ PER UNO SCOPO SPECIFICO E DI NON VIOLAZIONE.

Indice analitico

C

campione e sfidante 20
campione/sfidante 26, 27, 28, 29, 31, 32
cartella
 definizioni risorse 7
cluster di server 2, 6
Cognos
 campione/sfidante 32
collaborazione 1
confronto
 modelli 20, 25, 26

D

definizioni risorse
 definizioni server 7
definizioni server 7
 aggiunta 7
 IBM SPSS Modeler 7
 modifica 8
distribuzione 2

F

fasi
 IBM SPSS Modeler 9, 12, 13, 14, 15,
 19, 23
file
 IBM SPSS Modeler 9, 12, 13, 14, 15,
 19, 23
file di dati
 campione/sfidante 31
flusso
 IBM SPSS Modeler 9, 12, 13, 14, 15,
 19, 23

I

IBM SPSS Collaboration and Deployment
Services Deployment Manager 2, 4
IBM SPSS Collaboration and Deployment
Services Deployment Portal 2, 5
IBM SPSS Collaboration and Deployment
Services Repository 2, 3
IBM SPSS Modeler 2
 definizione server 7
 flusso 9, 12, 13, 14, 15, 19, 23
 server di esecuzione 2, 6, 7

L

lavori
 IBM SPSS Modeler 9, 12, 13, 14, 15,
 19, 23

M

modelli
 campione e sfidante 20
 campione/sfidante 25
 confronto 20, 25
 IBM SPSS Modeler 20, 25
 monitoraggio 20, 25
 valutazione 20, 25
modelli campione 20, 25, 26, 28, 29
modelli sfidante 20, 25, 26, 27, 28
 origine dati 28

N

nome della fase di lavoro
 campione/sfidante 26

O

origine dati ODBC
 campione/sfidante 32

P

parametri 19
preferenze utente 5

S

script
 campione e sfidante 20
 campione/sfidante 25
 IBM SPSS Modeler 20, 25
server di esecuzione 6
 IBM SPSS Modeler 2, 6
servizio di calcolo del punteggio 23
supernodi 19

V

vista dati analitici
 campione/sfidante 31



Stampato in Italia